

# Lawrence Berkeley National Laboratory

## Environ Genomics & Systems Bio

### Title

Draft Genome Sequence of Pedococcus sp. Strain 5OH\_020, Isolated from California Grassland Soil

### Permalink

<https://escholarship.org/uc/item/9p69j5mt>

### Journal

Microbiology Resource Announcements, 12(6)

### ISSN

2576-098X

### Authors

Green, Myka A  
Alvarez-Aponte, Zoila I  
Trotter, Valentine V  
et al.

### Publication Date

2023-06-20

### DOI

10.1128/mra.00025-23





### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



# Draft Genome Sequence of *Pedococcus* sp. Strain 5OH\_020, Isolated from California Grassland Soil

 Myka A. Green,<sup>a</sup>  Zoila I. Alvarez-Aponte,<sup>a</sup>  Valentine V. Trotter,<sup>b</sup>  Michiko E. Taga<sup>a</sup>

<sup>a</sup>Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, California, USA

<sup>b</sup>Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

**ABSTRACT** The draft genome sequence of the soil bacterium *Pedococcus* sp. strain 5OH\_020, isolated on a natural cobalamin analog, comprises 4.4 Mbp, with 4,108 protein-coding genes. Its genome encodes cobalamin-dependent enzymes, including methionine synthase and class II ribonucleotide reductase. Taxonomic analysis suggests that it is a novel species within the genus *Pedococcus*.

*Pedococcus* sp. strain 5OH\_020 was isolated from ungrazed grassland topsoil at the Hopland Research and Extension Center (39.00 N, 123.08 W) by limiting dilution (1) in methionine-dropout VL60 medium containing 0.1 g/L each of xylan, glucose, xylose, and *N*-acetyl glucosamine (2) amended with 10 nM 5-hydroxybenzimidazolyl cobamide ([5-OHBza]Cba) that we synthesized according to reference 3. [5-OHBza]Cba is a commercially unavailable cobamide cofactor typically biosynthesized by methanogens and found in diverse environments, such as animal guts and contaminated groundwater (4–6). To our knowledge, this is the first report of a bacterium isolated on a cobamide other than vitamin B<sub>12</sub> (cobalamin).

Strain 5OH\_020 was cultured at 28°C in the isolation medium, or on agar plates with the same medium at 2× concentration. The isolate is aerobic, Gram positive (7), and cocci-shaped (Fig. 1, inset), forming circular, cream-colored colonies after 12 days and growing only when [5-OHBza]Cba or cobalamin is provided. We Sanger sequenced 16S rRNA amplicons using the primers 27F and 1492R. After noting the isolate's low similarity to its closest NCBI BLAST database pairwise matches (<98.65%) (8), we performed whole-genome sequencing.

DNA was extracted from a pure liquid culture using the DNeasy blood and tissue kit (Qiagen) protocol for Gram-positive bacteria. Subsequent library preparation and sequencing steps were carried out by Novogene Bioinformatics Technology Co., Ltd. (Beijing, China). Library construction was performed using the NEBNext DNA library prep kit, following the manufacturer's instructions. After end repair, dA-tailing, and further ligation with NEBNext adapters, the DNA was sequenced using the NovaSeq 6000 platform (Illumina Inc., San Diego, CA, USA) to generate 150-bp paired-end reads, yielding 9,600,674 total reads after a raw data filtering step using FASTP v0.32.2 (9).

Data processing and analysis were performed using the Department of Energy Systems Biology Knowledgebase (10). The raw reads were trimmed and quality controlled using Trimmomatic v0.36 (11). The genome was assembled *de novo* via SPAdes v3.15.3 (12) and quality checked using QUAST v4.4 (13), generating the assembly statistics in Table 1. The genome was annotated and reannotated using Prokka v1.14.5 (14) and distilled using DRAM v0.1.2 (15). The 4.4-Mbp genome was quality checked using CheckM v1.0.18 (16). Default parameters were used for all software.

The genome was classified as belonging to the genus *Pedococcus*, with no further classification at the species level using GTDB-Tk v1.7.0 (17). The amino acid identity with *Pedococcus* sp. strain Soil748 was 95.61%. Using FastANI v1.33, an average nucleotide identity of 81.17% was indicated with the genus type strain *Pedococcus dokdonensis* (18). *Pedococcus* sp. 5OH\_020 forms an independent branch within the *Pedococcus* clade (Fig. 1) (19).

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

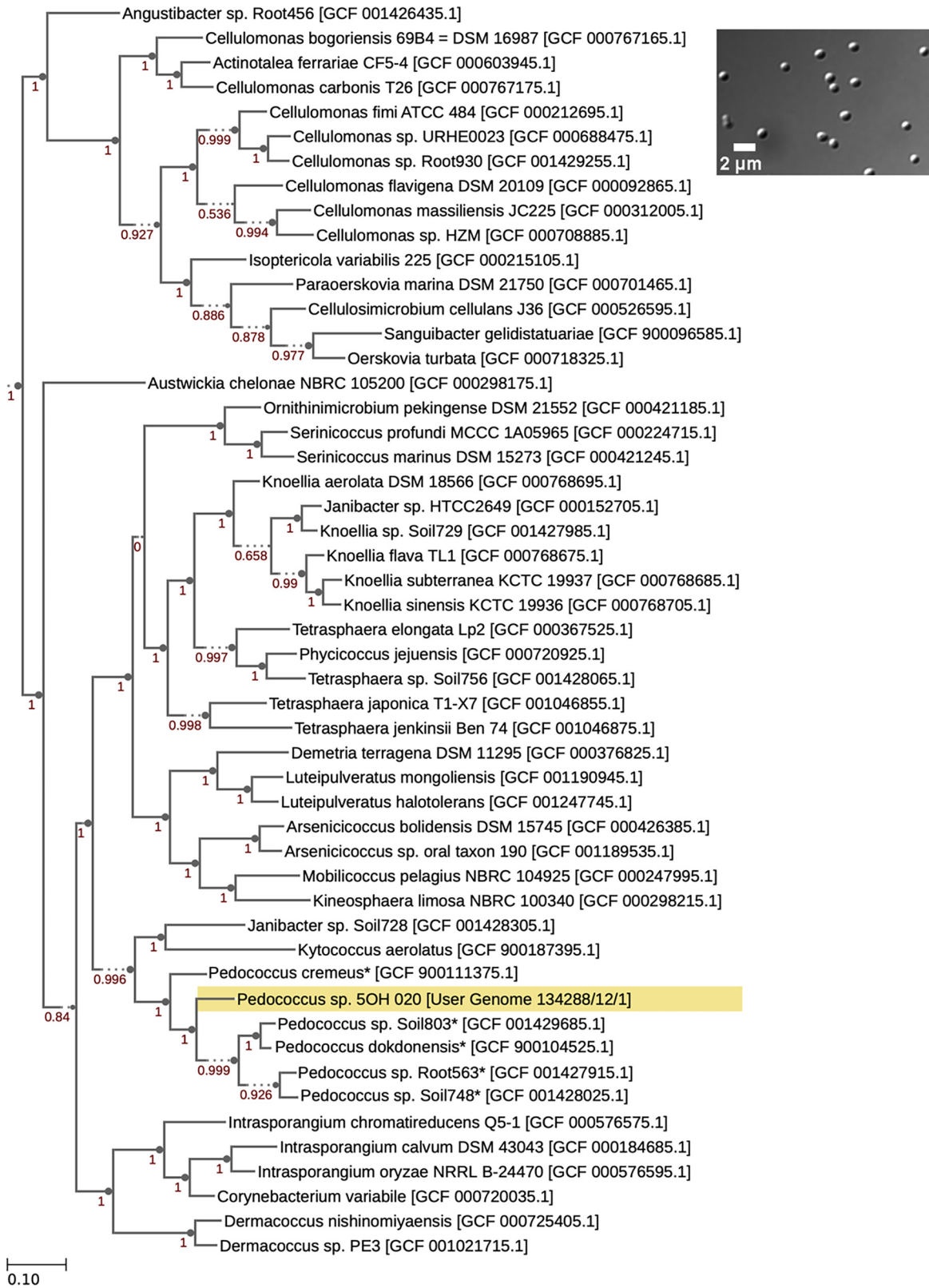
**Copyright** © 2023 Green et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Michiko E. Taga, taga@berkeley.edu.

The authors declare no conflict of interest.

**Received** 15 January 2023

**Accepted** 20 April 2023



**FIG 1** Morphology and phylogeny of *Pedococcus* sp. 5OH\_020. A phylogenetic tree was generated using the Insert Genome into Species Tree v2.2.0 (19) application, which uses the FastTree phylogenetic inference method. Parameters were set to include 50 neighboring species representative of the class Actinomycetia. Strain 5OH\_020 is highlighted in yellow. Asterisks indicate that the genus name *Phycococcus* was changed to *Pedococcus* to match current taxonomic classifications according to the NCBI database. Numbers in red represent bootstrap values (1,000 replicates). (Inset) Differential interference contrast (DIC) microscopy of *Pedococcus* sp. 5OH\_020 cells at  $\times 100$  magnification using a Zeiss Axiolmager DIF M1 microscope. Image captured using a Hamamatsu Orca 03 camera.

**TABLE 1** Assembly statistics generated using QUAST v4.4

Assembly statistic	Result
No. of contigs	33
Largest contig (bp)	783,609
Total length (bp)	4,427,096
Coverage (×)	325
$N_{50}$ (bp)	428,675
$N_{75}$ (bp)	278,038
$L_{50}$	5
$L_{75}$	8
G+C content (%)	69.75

*Pedococcus* sp. 5OH\_020 belongs to the phylum *Actinobacteria*, which is renowned for its production of clinically useful natural products (20). Among the 4,108 annotated protein-coding genes, two natural product biosynthetic clusters were detected using antiSMASH v6.0 (21). The genome also encodes enzymes predicted to degrade arabinan, xyloglucan, and amorphous cellulose, suggesting that strain 5OH\_020 may be capable of degrading plant matter in soil.

The genome encodes two cobalamin-dependent enzymes: methionine synthase and class II ribonucleotide reductase. No cobamide biosynthesis genes were present, consistent with the cobamide dependence observed in culture. Thus, strain 5OH\_020 likely relies on other microbes in the soil for cobamides.

**Data availability.** The draft genome sequence of *Pedococcus* sp. 5OH\_020 was deposited at DDBJ/ENA/GenBank under the accession number [JAPYZV000000000](https://doi.org/10.1093/jks.0.059774-0). The version described in this paper is version [JAPYZV010000000](https://doi.org/10.1093/jks.0.059774-0). The NCBI SRA accession number for the raw reads is [SRR22578002](https://doi.org/10.1093/jks.0.059774-0). The BioProject and BioSample accession numbers for this project are [PRJNA891762](https://doi.org/10.1093/jks.0.059774-0) and [SAMN31357841](https://doi.org/10.1093/jks.0.059774-0), respectively.

## ACKNOWLEDGMENTS

We thank Denise Schichnes and the RCNR Biological Imaging Facility at the University of California, Berkeley, for microscopy guidance, and Hans Carlson and Adam Deutschbauer for their advice on whole-genome sequencing. This research was supported by the Sponsored Projects for Undergraduate Research (SPUR) program at UC Berkeley (M.A.G.) and U.S. Department of Energy grant DE-SC0020155 to M.E.T. Z.I.A.-A. received support from a Grace Kase fellowship from the Kase/Tsujimoto Foundation and a GEM fellowship from the National GEM Consortium.

## REFERENCES

- Goodman AL, Kallstrom GK, Faith JJ, Reyes A, Moore A, Dantas G, Gordon JL. 2011. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci U S A* 108:6252–6257. <https://doi.org/10.1073/pnas.1102938108>.
- Sait M, Hugenholtz P, Janssen PH. 2002. Cultivation of globally distributed soil bacteria from phylogenetic lineages previously only detected in cultivation-independent surveys. *Environ Microbiol* 4:654–666. <https://doi.org/10.1046/j.1462-2920.2002.00352.x>.
- Mok KC, Hallberg ZF, Taga ME. 2022. Purification and detection of vitamin B<sub>12</sub> analogs. *Methods Enzymol* 668:61–85. <https://doi.org/10.1016/bs.mie.2021.11.023>.
- Hallberg ZF, Seth EC, Thevasundaram K, Taga ME. 2022. Comparative analysis of corrinoid profiles across host-associated and environmental samples. *Biochemistry* 61:2791–2796. <https://doi.org/10.1021/acs.biochem.2c00367>.
- Men Y, Seth EC, Yi S, Crofts TS, Allen RH, Taga ME, Alvarez-Cohen L. 2015. Identification of specific corrinoids reveals corrinoid modification in dechlorinating microbial communities. *Environ Microbiol* 17:4873–4884. <https://doi.org/10.1111/1462-2920.12500>.
- Shelton AN, Seth EC, Mok KC, Han AW, Jackson SN, Haft DR, Taga ME. 2019. Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. *ISME J* 13:789–804. <https://doi.org/10.1038/s41396-018-0304-9>.
- Popescu A, Doyle RJ. 1996. The Gram stain after more than a century. *Biotech Histochem* 71:145–151. <https://doi.org/10.3109/10520299609117151>.
- Kim M, Oh H-S, Park S-C, Chun J. 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 64:346–351. <https://doi.org/10.1099/ijs.0.059774-0>.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, 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>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Bankovich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm

- and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
13. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUILT: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
  14. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
  15. Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL, Solden LM, Liu P, Narrowe AB, Rodríguez-Ramos J, Bolduc B, Gazitúa MC, Daly RA, Smith GJ, Vik DR, Pope PB, Sullivan MB, Roux S, Wrighton KC. 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res* 48:8883–8900. <https://doi.org/10.1093/nar/gkaa621>.
  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. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>.
  18. Jain C, Rodríguez-Ramos 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>.
  19. 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>.
  20. Nouioui I, Carro L, Lopez-García M, Meier-Kolthoff JP, Woyke T, Kyrpides NC, Pukall R, Klenk HP, Goodfellow M, Göker M. 2018. Genome-based taxonomic classification of the phylum Actinobacteria. *Front Microbiol* 9:2007. <https://doi.org/10.3389/fmicb.2018.02007>.
  21. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. 2021. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res* 49:W29–W35. <https://doi.org/10.1093/nar/gkab335>.