



Complete Genome Sequence of *Starkeya* sp. Strain ORNL1, a Soil Alphaproteobacterium Isolated from the Rhizosphere of *Populus deltoides*

 Mircea Podar,^a Joel Turner,^a Leah H. Burdick,^a Dale A. Pelletier^a

^aBiosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

ABSTRACT *Starkeya* sp. strain ORNL1 is an alphaproteobacterium isolated from the rhizosphere of an Eastern cottonwood tree. *Starkeya* spp. are physiologically versatile, using a wide range of nutritional and energetic resources and serving important ecological roles in carbon and sulfur cycling. The 6.3-Mb chromosome of *Starkeya* sp. strain ORNL1 was completely sequenced and will help in understanding nutrient cycles.

Starkeya is a genus of alphaproteobacteria classified into the family *Xanthobacteraceae*, most closely related to *Ancylobacter* (1, 2). Its type species, *Starkeya novella*, was isolated from soil in the mid-1930s and was originally classified as *Thiobacillus* (1, 3). A second formally described species, *S. koreensis*, was isolated from rice straw (4). *S. novella* has been studied extensively, being able to grow as a facultative chemoautotroph, both consuming and producing CO₂ (5–7), utilize a variety of C₁ substrates (methanol, formate, formamide) (8, 9), and oxidize sulfur compounds (9, 10).

Here, we report the complete genome sequence of a putative novel species of *Starkeya*, strain ORNL1, isolated from the microbial rhizosphere of an Eastern cottonwood tree (*Populus deltoides*) in Oak Ridge, Tennessee. We used flow cytometry to deposit single bacterial cells from a microbially enriched rhizosphere fraction onto Reasoner's 2A (R2A) agar (11, 12). Colonies that developed upon incubation at 28°C were identified taxonomically by small-subunit (SSU) rRNA gene amplification and direct Sanger sequencing using the universal primers 27F and 1492R (13). Default parameters were used for all software used for sequence data analysis unless otherwise specified. Sequences were assembled and manually trimmed based on quality in Geneious R11 (14), and close relatives were identified by MegaBLAST (15) against the NCBI rRNA database. The sequences and their top relatives were aligned with ClustalW 2.1 (16) in Geneious, and the alignment was manually edited to remove heterogeneous ends. One colony was identified as representing a *Starkeya* sp., sharing 98% pairwise sequence identity with the rRNA sequences of *S. novella* and *S. koreensis*. Phylogenetic analysis using FastTree 2.1.11 (17) indicated that the novel isolate, designated *Starkeya* sp. strain ORNL1, is most closely related to *S. novella* (Fig. 1).

For PacBio genomic sequencing, *Starkeya* sp. strain ORNL1 was grown in liquid R2A medium at 30°C for 2 days. Genomic DNA was purified using the Quick-DNA fungal/bacterial miniprep kit (Zymo Research) and fragmented to an average size of 10 kb using g-TUBEs (Covaris, Woburn, MA). A sequencing library was constructed using a SMRTbell template prep kit 1.0 (Pacific Biosciences, Menlo Park, CA) and sequenced on a Pacific Biosciences Sequel instrument. The sequences were filtered based on quality and assembled using the software HGAP4 in PacBio SMRTLink 7.0, using a target genome size of 6 Mbp, a minimum subread length of 500, a minimum concordance of 85%, and a seed coverage of 30-fold, with the “aggressive” option on. In all, 64,603 filtered subreads (N_{50} , 10,183 nucleotides [nt]) were assembled into three polished

Citation Podar M, Turner J, Burdick LH, Pelletier DA. 2020. Complete genome sequence of *Starkeya* sp. strain ORNL1, a soil alphaproteobacterium isolated from the rhizosphere of *Populus deltoides*. Microbiol Resour Announc 9:e00644-20. <https://doi.org/10.1128/MRA.00644-20>.

Editor J. Cameron Thrash, University of Southern California

Copyright © 2020 Podar 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 Mircea Podar, podarm@ornl.gov.

Received 2 June 2020

Accepted 16 June 2020

Published 2 July 2020

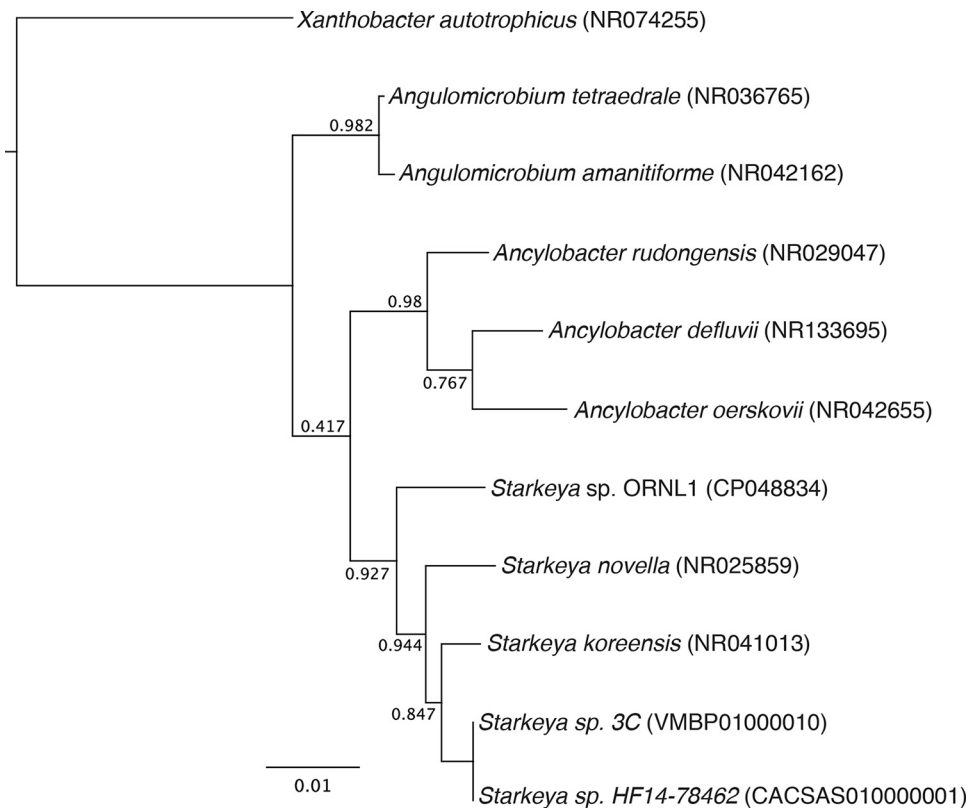


FIG 1 Phylogenetic tree (FastTree) of *Starkeya* sp. strain ORNL1 and related bacteria based on SSU rRNA genes. Reference sequences were obtained from GenBank; accession numbers are in parentheses. The alignment was edited by removing heterogeneous-length ends. The tree was rooted with *Xanthobacter*. Numbers at the nodes indicate support values.

contigs totaling 6,326,767 nt, with a mean coverage of 350-fold and a G+C content of 66%. We predicted the genes using Prokka (18) and performed *de novo* assembly in Geneious 11 (14), together with PacBio spanning sequence reads. All three contigs were self-assembled, and based on gene calls, we identified overlapping assembly ends that enabled closing of the genome as a single, circular chromosome of 6,286,188 bp. Gene prediction and functional annotation were performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) 4.8 (19), which identified 5,826 protein coding sequences, 48 tRNAs, 2 rRNA operons, and 13 other noncoding or regulatory RNAs (ncRNAs). We also generated a metabolic model in KBase (20), accessible at <https://narrative.kbase.us/narrative/55377>, together with its RAST annotation. The average nucleotide identity (ANI) of 83% relative to *S. novella*, calculated with FastANI 0.1.2 (21), suggests that *Starkeya* sp. strain ORNL1 may represent a novel species, provisionally referred to as *Starkeya rhizosphaerae* ORNL1. Its genome is significantly larger than that of *S. novella* (4.7 Mbp) and will aid in understanding adaptation to the rhizosphere microbiome community, as well as uncovering potential broader metabolic versatility in this group of bacteria.

Data availability. The *Starkeya* sp. strain ORNL1 genome sequence has been deposited in GenBank under the accession number [CP048834](https://www.ncbi.nlm.nih.gov/nuccore/CP048834). The version described in this paper is the first version, [CP048834.1](https://www.ncbi.nlm.nih.gov/nuccore/CP048834.1). The PacBio reads have been deposited in the SRA under the accession number [SRX7858655](https://www.ncbi.nlm.nih.gov/sra/SRX7858655).

ACKNOWLEDGMENTS

We thank the Genomic Resource Center at the University of Maryland School of Medicine for PacBio sequencing.

This research was funded by the U.S. DOE Office of Biological and Environmental

Research, Genomic Science Program, 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.

REFERENCES

- Kelly DP, McDonald IR, Wood AP. 2000. Proposal for the reclassification of *Thiobacillus novellus* as *Starkeya novella* gen. nov., comb. nov., in the alpha-subclass of the Proteobacteria. *Int J Syst Evol Microbiol* 50:1797–1802. <https://doi.org/10.1099/00207713-50-5-1797>.
- Kappler U, Davenport K, Beatson S, Lucas S, Lapidus A, Copeland A, Berry KW, Glavina Del Rio T, Hammon N, Dalin E, Tice H, Pitluck S, Richardson P, Bruce D, Goodwin LA, Han C, Tapia R, Detter JC, Chang Y-J, Jeffries CD, Land M, Hauser L, Kyrpides NC, Goker M, Ivanova N, Klenk H-P, Woyke T. 2012. Complete genome sequence of the facultatively chemolithoautotrophic and methylotrophic alpha proteobacterium *Starkeya novella* type strain (ATCC 8093^T). *Stand Genomic Sci* 7:44–58. <https://doi.org/10.4056/sogs.3006378>.
- Starkey RL. 1934. Isolation of some bacteria which oxidize thiosulfate. *Soil Sci* 39:197–220.
- Im W-T, Aslam Z, Lee M, Ten LN, Yang D-C, Lee S-T. 2006. *Starkeya koreensis* sp. nov., isolated from rice straw. *Int J Syst Evol Microbiol* 56:2409–2414. <https://doi.org/10.1099/ijs.0.64093-0>.
- McCarthy JT, Charles AM. 1974. CO₂ fixation by the facultative autotroph *Thiobacillus novellus* during autotrophy–heterotrophy interconversions. *Can J Microbiol* 20:1577–1584. <https://doi.org/10.1139/m74-244>.
- Perez RC, Matin A. 1982. Carbon dioxide assimilation by *Thiobacillus novellus* under nutrient-limited mixotrophic conditions. *J Bacteriol* 150:46–51. <https://doi.org/10.1128/JB.150.1.46-51.1982>.
- Santer M, Boyer J, Santer U. 1959. *Thiobacillus novellus*. I. Growth on organic and inorganic media. *J Bacteriol* 78:197–202. <https://doi.org/10.1128/JB.78.2.197-202.1959>.
- Chandra TS, Shethna YI. 1977. Oxalate, formate, formamide, and methanol metabolism in *Thiobacillus novellus*. *J Bacteriol* 131:389–398. <https://doi.org/10.1128/JB.131.2.389-398.1977>.
- Kappler U, Nouwens AS. 2013. Metabolic adaptation and trophic strategies of soil bacteria-C1- metabolism and sulfur chemolithotrophy in *Starkeya novella*. *Front Microbiol* 4:304. <https://doi.org/10.3389/fmicb.2013.00304>.
- Kappler U, Friedrich CG, Truper HG, Dahl C. 2001. Evidence for two pathways of thiosulfate oxidation in *Starkeya novella* (formerly *Thiobacillus novellus*). *Arch Microbiol* 175:102–111. <https://doi.org/10.1007/s002030000241>.
- Podar M, Turner J, Burdick LH, Pelletier DA. 2019. Complete genome sequence of *Terriglobus albidus* strain ORNL, an acidobacterium isolated from the *Populus deltoides* rhizosphere. *Microbiol Resour Announc* 8:e01065-19. <https://doi.org/10.1128/MRA.01065-19>.
- Utturkar SM, Cude WN, Robeson MS, Jr, Yang ZK, Klingeman DM, Land ML, Allman SL, Lu T-Y, Brown SD, Schadt CW, Podar M, Doktycz MJ, Pelletier DA. 2016. Enrichment of root endophytic bacteria from *Populus deltoides* and single-cell-genomics analysis. *Appl Environ Microbiol* 82:5698–5708. <https://doi.org/10.1128/AEM.01285-16>.
- Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ. 2008. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Appl Environ Microbiol* 74:2461–2470. <https://doi.org/10.1128/AEM.02272-07>.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.
- Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schaffer AA. 2008. Database indexing for production MegaBLAST searches. *Bioinformatics* 24:1757–1764. <https://doi.org/10.1093/bioinformatics/btn322>.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948. <https://doi.org/10.1093/bioinformatics/btm404>.
- 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>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Tatusova T, DiCuccio M, Badretdin A, Chetverin 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>.
- 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>.
- 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>.