Microbiology Resource Announcements

| Environmental Microbiology | Announcement

Complete genome sequence of *Roseinatronobacter* sp. strain S2, a chemolithoheterotroph isolated from a pH 12 serpentinizing system

Leah R. Trutschel,[1] Annette R. Rowe,[1] Joshua D. Sackett[1]

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT Here, we report the complete genome sequence for *Roseinatronobacter* sp. S2, a sulfur-oxidizing heterotroph isolated from a serpentinizing system in Northern California. The S2 genome is 4.4 Mb and contains 4,570 protein-encoding genes. This organism contains the genes necessary for sulfur species oxidation and complete ethylmalonyl and pentose phosphate pathways.

KEYWORDS serpentinization, alkaliphile, sulfur oxidation, chemolithoheterotroph, Ney Springs

Serpentinizing systems are hyperalkaline (pH >10) environments characterized by the presence of abiogenic hydrogen and methane (1–5). Ney Springs is a pH 12–12.5 marine-like spring characterized by high amounts of sulfide, ammonia, and methane (6). Many abundant microorganisms in this spring are putative sulfur oxidizers belonging to the *Paracoccaceae* (formerly *Rhodobacteraceae)* (6). *Roseinatronobacter* sp. S2 is one of the first sulfur-oxidizing isolates obtained from a serpentinizing system (6); its completed genome provides insight into the adaptations and metabolism of alkaliphilic chemolithoheterotrophs.

S2 was isolated from Ney Springs in Mt. Shasta, California, on minimal media agar plates containing 20 mM polysulfide and 10 mM acetate (6). S2 was cultivated aerobically in liquid minimal media containing 20 mM thiosulfate and 10 mM acetate. Detailed media instructions are found here: dx.doi.org/10.17504/protocols.io.bqjgmujw. S2 was incubated at room temperature for 5 days to achieve approximate maximum turbidity as determined by a previous growth curve (6). DNA was extracted using a Qiagen DNeasy PowerSoil Kit and quantified using a Qubit fluorometer (ThermoFisher Scientific, USA). All sequencing was done from a single DNA prep. Nanopore libraries were prepped with the Native Barcoding 24 V14 Kit (Oxford Nanopore Technologies, Oxford, UK) and sequenced using a R10.4.1 flow cell (FLO-MIN114) under high-accuracy mode (280 bp/s) with a MinION-sequencing device. Basecalls were made with Guppy v.6.4.6, and reads with quality scores <7 were removed (7). Illumina library prep and sequencing were conducted at SeqCenter LLC (Pittsburgh, PA, USA). Briefly, libraries were prepared with the Illumina DNA Prep Kit, barcoded with 10 bp unique dual indices, and sequenced on an Illumina NovaSeq (2 × 150 sequencing). Demultiplexing, quality control, and adapter trimming were performed with bcl-convert (v.4.0.3). Nanopore sequences >2,000 bp were quality filtered with Filtlong (v.0.2.1) (8), and the resulting worst 10% of read bases were removed. Filtered long reads were assembled with Flye (v.2.9.1) (9). Illumina reads were quality filtered with FastQC (v.0.12.1) (10) with all reads passing with a quality score >Q30. Short reads were aligned to the assembly with Burrows-Wheeler Aligner (v.0.7.17) (11), and the assembly was polished with Pilon (v.1.24, -fix all) (12). Four rounds of polishing were performed. Quality assessment and genome statistics were

Editor Julia A. Maresca, University of Delaware College of Engineering, Newark, Delaware, USA

Address correspondence to Annette R. Rowe, annette.rowe@uc.edu.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 4 April 2023
Accepted 13 July 2023
Published 16 August 2023

TABLE 1 Summary of genome statistics

Parameter	Finding
Illumina reads accession no.	SRR23870722
MinION reads accession no.	SRR23870721
Genome assembly accession no.	CP121113, CP121114, CP121115, CP121116, CP121117, CP121118
Assembly N_{50} (bp)[a]	3,479,920
Nanopore N_{50} (bp)[b]	11,120
Assembly size (bp)	4,445,845
Nanopore read length (bp)[b]	686,780,239
Illumina read count (seqs)	2,475,882
G + C content (%)	59.5
Estimated genome completeness (%)[c]	98.47
Estimated genome contamination (%)[c]	1.21
Estimated Nanopore coverage of largest contig (Genome) (×)[b]	155
Estimated Illumina coverage (×)	152
No. of contigs	6
Contig_1 length (chromosome) (bp)	3,471,855
Contig_3 length (bp)	611,552
Contig_5 length (bp)	166,770
Contig_4 length (bp)	136,430
Contig_6 length (bp)	33,793
Contig_2 length (bp)	15,307
No. of protein-coding genes	4,570
No. of tRNAs	47
No. of rRNA operons	3

[a]Determined with Quast (13).
[b]Determined with Flye (9).
[c]Determined with CheckM (14).

determined using QUAST (v.4.4) (13) and CheckM (v.1.0.18) (14) within the KBase wrapper (15). Taxonomic classification was done with GTDB-tK (V. 2.2.5) (16). The genome was annotated with the NCBI Prokaryotic Genome Annotation Pipeline v6.5 (17). Default settings were used for all programs unless noted.

The final assembly size for the genome was 4,445,845 bp, and it has a GC content of 59.5%. The assembly was composed of six circular contigs, potentially chromids or plasmids, as often seen in members of the *Paracoccaceae* (18). GTDB-Tk classified the assembly as *Roseinatronobacter* with an average nucleotide identity of 84.66% with the nearest neighbor *Roseinatronobacter* sp. 017510335 (alignment fraction = 0.747). S2 is predicted to contain the ethylmalonyl pathway, which has been used as a way to utilize acetate in bacteria lacking isocitrate lyase (19). S2 contains SoxBCDYZ, Sqr, and FccB and is capable of thiosulfate oxidation to sulfate *in vitro* (6).

ACKNOWLEDGMENTS

This work was funded by NSF-EAR LowTemp Geochemistry Geobiology award 2025687 and NASA-Roses Exobiology Program grant number 80NSSC21K0482, awarded to A.R.R., and NSF Postdoctoral Research Fellowship NSF-OCE 2126677 awarded to J.D.S.

AUTHOR AFFILIATION

[1]Department of Biological Sciences, University of Cincinnati, Cincinnati, Ohio, USA

AUTHOR ORCIDs

Leah R. Trutschel http://orcid.org/0000-0002-0596-6977

FUNDING

Funder	Grant(s)	Author(s)
National Science Foundation (NSF)	NSF-OCE 2126677	Joshua Sackett
National Science Foundation (NSF)	NSF-EAR 2025687	Annette Ruth Rowe
National Aeronautics and Space Administration (NASA)	80NSSC21K0482	Annette Ruth Rowe

DATA AVAILABILITY

The raw sequencing data were submitted to the NCBI Sequence Read Archive, and the genome assembly was submitted to GenBank under the accession numbers listed in Table 1.

REFERENCES

1. Schrenk MO, Brazelton WJ, Lang SQ. 2013. Serpentinization, carbon, and deep life. Reviews in Mineralogy and Geochemistry 75:575–606. https://doi.org/10.2138/rmg.2013.75.18
2. Mccollom TM, Seewald JS. 2013. Serpentinites, hydrogen, and life:129–134.
3. Sabuda MC, Brazelton WJ, Putman LI, McCollom TM, Hoehler TM, Kubo MDY, Cardace D, Schrenk MO. 2020. A dynamic microbial sulfur cycle in a serpentinizing continental ophiolite. Environ Microbiol 22:2329–2345. https://doi.org/10.1111/1462-2920.15006
4. Brazelton WJ, Schrenk MO, Kelley DS, Baross JA. 2006. Methane- and sulfur-metabolizing microbial communities dominate the lost city hydrothermal field ecosystem. Appl Environ Microbiol 72:6257–6270. https://doi.org/10.1128/AEM.00574-06
5. Glombitza C, Putman LI, Rempfert KR, Kubo MD, Schrenk MO, Templeton AS, Hoehler TM. 2021. Active microbial sulfate reduction in fluids of serpentinizing peridotites of the continental subsurface. Commun Earth Environ 2:1–9. https://doi.org/10.1038/s43247-021-00157-z
6. Trutschel LR, Chadwick GL, Kruger B, Blank JG, Brazelton WJ, Dart ER, Rowe AR. 2022. Investigation of microbial metabolisms in an extremely high pH marine-like terrestrial serpentinizing system: ney springs. Sci Total Environ 836:155492. https://doi.org/10.1016/j.scitotenv.2022.155492
7. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. Sequence analysis nanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149
8. Wick R. 2021. Filtlong. Github repository
9. 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
10. Andrews S. 2015. FastQC: a quality control tool for high throughput sequence data [Online].
11. Li H, Durbin R. 2010. Fast and accurate long-read alignment with burrows-wheeler transform. Bioinformatics 26:589–595. https://doi.org/10.1093/bioinformatics/btp698
12. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963
13. 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
14. 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
15. 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, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. Kbase: the United States department of energy systems biology knowledgebase. Nat Biotechnol 36:566–569. https://doi.org/10.1038/nbt.4163
16. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH, Hancock J. 2019. GTDB-TK: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848
17. Tatusova T, DiCuccio M, Badretdin A, Chetvernin 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
18. Petersen J, Frank O, Göker M, Pradella S. 2013. Extrachromosomal, extraordinary and essential--the plasmids of the roseobacter clade. Appl Microbiol Biotechnol 97:2805–2815. https://doi.org/10.1007/s00253-013-4746-8
19. Erb TJ, Berg IA, Brecht V, Müller M, Fuchs G, Alber BE. 2007. Synthesis of C_5-dicarboxylic acids from C_2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. Proc Natl Acad Sci U S A 104:10631–10636. https://doi.org/10.1073/pnas.0702791104