

Complete genome sequence of *Thioclava* strain BDS19, isolated from coastal marsh sediment

Biagio DiSalvo,^{1,2} Blair G. Paul,² Zoe G. Cardon¹

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Salt marsh sediments host intensive sulfur cycling. Genome sequences from culturable sulfur-oxidizing microbes are required for targeted engineering to test *sox* gene function. We report the complete genome sequence (3,410,538 bp, 59.6 mol% G + C content) of putative sulfur oxidizer *Thioclava* strain BDS19, isolated from marsh sediment on Cape Cod, USA.

KEYWORDS *Thioclava*, genomes, sulfur oxidation, sulfur oxidizer, salt marsh

Sulfur-cycling microbes are integral members of coastal marsh ecosystems (1). Heterotrophic sulfate reducers transform ocean-derived sulfates into energy-containing forms such as sulfide and thiosulfate, which sulfur oxidizers can then use to fuel growth. The forms of oxidized and reduced sulfur present in marsh sediments are important, as, for example, sulfide can be toxic to aerobic respiration in diverse plants (2). Though microbes involved in linked sulfur- and carbon-cycling have been studied (3, 4), genetic tools are needed to characterize the functionality of genes underlying these processes. Complete genomes provide necessary maps to identify target genes of interest and enable the design of sequence-specific engineering tools (5).

An isolate of *Thioclava* was obtained from pink sediment mats collected in March 2025 at *Sporobolus alterniflorus*-dominated Little Sippewissett Marsh [41°34'34.0"N, 70°38'18.8"W], Falmouth, MA. Bacteria in this genus are gram-negative, rod-shaped *Alphaproteobacteria* with ecological functions including carbon dioxide fixation, crude oil degradation, and sulfur oxidation (6, 7). They can be halotolerant, facultative anaerobes, and mixotrophic (8). Here, we report the complete genome sequence of *Thioclava* strain BDS19 as a first step toward engineering to understand regulation and function of sulfur cycling-related genes.

Briefly, a suspension of sediment was plated onto Luria-Bertani (LB) agar plates and incubated at 20–22°C for 1 week. Individual pale orange colonies were T-struck to obtain a pure culture. Strain BDS19 was stored in 25% (w/v) glycerol at –80°C. For whole genome sequencing, cells were harvested from 1-week-old plates, inoculated into LB broth, and incubated in a shaker incubator at 30°C at 193 rpm overnight. Genomic DNA was extracted using the NEB Monarch Genomic DNA Extraction and Purification Kit (#T3010S, NEB, USA).

Libraries were generated by Plasmidsaurus (Louisville, KY, USA) using an Ultra-long Sequencing Kit V14 by Oxford Nanopore Technology (ONT) (Lexington, MA, USA) and sequenced on PromethION R10.4.1 flow cells. Default parameters were used for all software unless otherwise specified. ONT's Dorado (10.4.1) was used for basecalling and SPRI beads (Omega Bio-Tek, Georgia, USA) for size selection. A total of 267,568,203 bp with 74,581 reads were sequenced, and N50 was 9,100 bp. Filtlong (v0.2.1) (9) filtered out the lowest 5% of FASTQ reads by quality and downsampled the reads to 250 Mb to create a rough assembly using Miniasm (v0.3) (10). The information generated by Miniasm was used to further downsample to gather reads at up to ~100× coverage,

Editor Elinne Becket, California State University San Marcos, San Marcos, California, USA

Address correspondence to Zoe G. Cardon, zcardon@mbi.edu.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 1 August 2025

Accepted 9 November 2025

Published 26 November 2025

Copyright © 2025 DiSalvo 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/).

with the parameter `fitlong-mean_q_weight 10` applied to remove additional low-quality reads. The resulting subsample reads were used by Flye v2.9.1 (11) to call the circularity of the chromosome and generate a final assembly, which was further processed with Medaka v1.8.0 (12).

Total genome size was 3,410,538 bp with a G + C content of 59.6 mol%. Estimated completeness was 98.8% with 0.15% contamination (CheckM [v.1.2.2]) (13). Mash (v.2.3) (14) identified its closest related species as *Thioclava indica* (NZ_AUBN01000001.1) with 90.4% identity. Genome annotation was performed by NCBI Prokaryotic Genome Annotation Pipeline (v.6.10) (15) and identified 3,297 putative genes, 48 tRNAs, 9 rRNAs, and 4 ncRNAs. Numerous sulfur-related genes were annotated, including a sulfur-oxidizing cluster, *soxABCXYZ*.

ACKNOWLEDGMENTS

This study was supported by Gordon and Betty Moore Foundation grant GBMF9347 to Z.G.C. and B.G.P.

AUTHOR AFFILIATIONS

¹The Ecosystems Center, Marine Biological Laboratory, Woods Hole, Massachusetts, USA

²Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, Massachusetts, USA

AUTHOR ORCIDs

Biagio DiSalvo  <http://orcid.org/0009-0002-4088-2776>

Blair G. Paul  <https://orcid.org/0000-0002-5738-5568>

Zoe G. Cardon  <http://orcid.org/0000-0001-8725-7842>

FUNDING

Funder	Grant(s)	Author(s)
Gordon and Betty Moore Foundation	9347	Zoe G. Cardon Blair G. Paul

AUTHOR CONTRIBUTIONS

Biagio DiSalvo, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Writing - original draft, Writing - review and editing | Blair G. Paul, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing - review and editing | Zoe G. Cardon, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing - review and editing

DATA AVAILABILITY

The draft genome has been submitted to GenBank, accession number [CP195919](#), under BioProject [PRJNA1285420](#) and the raw sequence file is available under SRA accession number [SRP604649](#).

REFERENCES

1. Pérez Castro S, Peredo EL, Mason OU, Vineis J, Bowen JL, Mortazavi B, Ganesh A, Ruff SE, Paul BG, Giblin AE, Cardon ZG. 2023. Diversity at single nucleotide to pangenome scales among sulfur cycling bacteria in salt marshes. *Appl Environ Microbiol* 89:e00988-23. <https://doi.org/10.1128/aem.00988-23>
2. Lamers LPM, Govers LL, Janssen ICJM, Geurts JJM, Van der Welle MEW, Van Katwijk MM, Van der Heide T, Roelofs JGM, Smolders AJP. 2013. Sulfide as a soil phytotoxin—a review. *Front Plant Sci* 4:268. <https://doi.org/10.3389/fpls.2013.00268>
3. Thomas F, Giblin AE, Cardon ZG, Sievert SM. 2014. Rhizosphere heterogeneity shapes abundance and activity of sulfur-oxidizing

- bacteria in vegetated salt marsh sediments. *Front Microbiol* 5:309. <https://doi.org/10.3389/fmicb.2014.00309>
4. Anantharaman K, Hausmann B, Jungbluth SP, Kantor RS, Lavy A, Warren LA, Rappé MS, Pester M, Loy A, Thomas BC, Banfield JF. 2018. Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle. *ISME J* 12:1715–1728. <https://doi.org/10.1038/s41396-018-0078-0>
 5. Song Y, He S, Jopkiewicz A, Setroikromo R, van Merkerk R, Quax WJ. 2022. Development and application of CRISPR-based genetic tools in *Bacillus* species and *Bacillus* phages. *J Appl Microbiol* 133:2280–2298. <https://doi.org/10.1111/jam.15704>
 6. Liu Y, Lai Q, Shao Z. 2017. A multilocus sequence analysis scheme for phylogeny of *Thioclava* bacteria and proposal of two novel species. *Front Microbiol* 8:1321. <https://doi.org/10.3389/fmicb.2017.01321>
 7. Yin X, Li X, Li Q, Wang B, Zheng L. 2023. Complete genome analysis reveals environmental adaptability of sulfur-oxidizing bacterium *Thioclava nitratreducens* M1-LQ-LJL-11 and symbiotic relationship with deep-sea hydrothermal vent *Chrysomallon squamiferum*. *Mar Genomics* 71:101058. <https://doi.org/10.1016/j.margen.2023.101058>
 8. Sorokin DY, Tourova TP, Spiridonova EM, Rainey FA, Muyzer G. 2005. *Thioclava pacifica* gen. nov., sp. nov., a novel facultatively autotrophic, marine, sulfur-oxidizing bacterium from a near-shore sulfidic hydrothermal area. *Int J Syst Evol Microbiol* 55:1069–1075. <https://doi.org/10.1099/ijs.0.63415-0>
 9. 2021. Fitlong. <https://github.com/rrwick/Fitlong>.
 10. Li Heng. 2016. Minimap and miniiasm: fast mapping and *de novo* assembly for noisy long sequences. *Bioinformatics* 32:2103–2110. <https://doi.org/10.1093/bioinformatics/btw152>
 11. 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>
 12. 2023. Medaka. <https://github.com/nanoporetech/medaka>.
 13. 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>
 14. Ondov BD, Starrett GJ, Sappington A, Kostic A, Koren S, Buck CB, Phillippy AM. 2019. Mash Screen: high-throughput sequence containment estimation for genome discovery. *Genome Biol* 20:232. <https://doi.org/10.1186/s13059-019-1841-x>
 15. 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>