

The complete genome sequences of *Thioglobus autotrophicus* strains EF2 and EF3, isolated from an oxycline in Effingham Inlet, British Columbia

Robert M. Morris,¹ Sayaka Mino^{1,2}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Here we provide the complete genome sequences of two chemoautotrophic isolates from the *Thioglobaceae* family of marine gamma-proteobacteria. The genomes were obtained from pure cultures that were initially isolated from Effingham Inlet in 2013 and revived from freezer stocks for whole genome sequencing in 2023.

KEYWORDS isolate, SUP05, *Thioglobaceae*, complete genome

Bacteria from the *Thioglobaceae* family (SUP05 clade) often dominate marine microbial communities in anoxic environments (1), including in marine oxygen minimum zones (2–7). Identifying the genetic potential for diverse *Thioglobaceae* to carry out denitrification is critical in understanding the marine nitrogen cycle.

Strains EF2 and EF3 were isolated from Effingham Inlet at 49.093 N 125.195 W in 2013 as previously described (8, 9). Briefly, isolates were cultured by diluting cells to extinction in 96-well plates containing filter sterilized seawater media (30 kDa) amended with 1 mM thiosulfate. Plates were incubated at 10°C and wells that were positive for growth after 3 weeks were screened by 16S rRNA analyses. Glycerol stocks (10%) were prepared and stored at –80°C. Cells were revived and grown aerobically in 1 L polycarbonate bottles at 13°C, then collected onto sterile, 47 mm, 0.2 µM polycarbonate filters (MilliporeSigma, Burlington, MA) after reaching maximum cell densities (~1 × 10⁶ cells/mL). The same DNA was used for Oxford Nanopore and Illumina sequencing. DNA was extracted using a QuickGene Tissue Kit and Mini-80 system (Holliston, MA) with minor modifications. Filters were submerged in TE buffer (pH 8.0) and incubated at –80°C for 30 min and at 95°C for 10 min. Lysates were then prepared according to the instructions. The lysates were processed through QuickGene Mini-80 cartridges by loading 750 µL at a time and then washing with WDT (wash buffer). DNA was eluted in water and stored at –80°C.

We used Illumina paired end (150 bp) and Oxford Nanopore sequences to construct assemblies. Software default parameters were used unless otherwise noted. Illumina sequencing libraries were constructed using a KAPA Library Preparation Kit and sequenced on a MiSeq instrument at the Northwest Genomics Center (Seattle, WA, USA). Adapter sequences, low-quality reads (<Q30), 1 bp at the tail end, reads shorter than 20 bp, and reads that had more than 10 numbers of N bases were removed using fastp v0.23.2 (10). Nanopore libraries were constructed using a Rapid V14 barcoding kit (SQK-RBK114.24) and sequenced using a MinION Mk1C and a R10.4.1 flow cell. Basecalling was performed using Guppy v6.4.6 in accurate mode. Reads were quality filtered using NanoFilt v2.8.0 (11) with a quality threshold of 10, a length threshold of 100 bp, and trimming of 75 bp from start of each read. Assemblies were obtained using Unicycler v0.4.8 (12) with all reads (Table 1) and produced single circular contigs. Average nucleotide identity (ANI) was calculated using the ani.rb script from the enve-omics toolkit (12) and indicates that EF2 and EF3 are strains related to *T. autotrophicus* strain

Editor J. Cameron Thrash, University of Southern California, USA

Address correspondence to Robert M. Morris, morrisrm@uw.edu.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 16 November 2023

Accepted 23 January 2024

Published 9 February 2024

Copyright © 2024 Morris and Mino. 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/).

TABLE 1 Data associated with sequencing and assembly *T. autotrophicus* EF2 and EF3

Strain	EF2	EF3
No. of raw reads		
Illumina	2,867,684	4,088,272
Nanopore	322,030	424,316
Nanopore raw read N ₅₀ (bp)	6,124	6,923
Assembled contig (bp)	1,497,563	1,442,386
Coverage	828 ×	1340 ×
Rotation	dnaA start	dnaA start
GC (%)	39.1	39.4
Genes	1,623	1,591
Protein-coding genes	1,573	1,545

EF1, with ANI% scores to strain EF1 of 97.8. Genomes were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (13).

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation OCE-1558483 and OCE-2201310 (R.M.) and the Japan Society for the Promotion of Science through Overseas Research Fellowships (S.M.).

We would also like to thank Dr. Richard Keil and the captain and crew of the RV Thompson for their assistance in collecting the samples used to isolate *T. autotrophicus* strains EF2 and EF3 during the University of Washington undergraduate senior thesis cruise of 2013 and Dr. Vega Shah for conducting the original cultivation work.

AUTHOR AFFILIATIONS

¹School of Oceanography, University of Washington, Seattle, Washington, USA

²Laboratory of Microbiology, Faculty of Fisheries Sciences, Hokkaido University, Minatocho, Hakodate, Japan

AUTHOR ORCIDS

Robert M. Morris  <http://orcid.org/0000-0001-9030-3790>

FUNDING

Funder	Grant(s)	Author(s)
National Science Foundation (NSF)	OCE-1558483	Robert M. Morris
National Science Foundation (NSF)	OCE-2201310	Robert M. Morris

AUTHOR CONTRIBUTIONS

Robert M. Morris, Conceptualization, Funding acquisition, Project administration, Writing – original draft | Sayaka Mino, Data curation, Formal analysis, Validation, Visualization, Writing – review and editing

DATA AVAILABILITY

Genome sequences were deposited in GenBank under BioSample, BioProject, Sequence Read Archive, and Genome assembly accession numbers: Strain EF2, [SAMN37983939](#), [PRJNA1032126](#), [SRR26639703](#), [SRR27349087](#), [SRR27349088](#), and [CP137632](#); Strain EF3, [SAMN37983940](#), [PRJNA1032127](#), [SRR26660965](#), [SRR27348713](#), and [CP137633](#).

REFERENCES

1. Morris RM, Spietz RL. 2022. The physiology and biogeochemistry of SUP05. *Ann Rev Mar Sci* 14:261–275. <https://doi.org/10.1146/annurev-marine-010419-010814>
2. Lavik G, Stührmann T, Brüchert V, Van der Plas A, Mohrholz V, Lam P, Mussmann M, Fuchs BM, Amann R, Lass U, Kuypers MMM. 2009. Detoxification of sulphidic African shelf waters by blooming chemolithotrophs. *Nature* 457:581–584. <https://doi.org/10.1038/nature07588>
3. Walsh DA, Zaikova E, Howes CG, Song YC, Wright JJ, Tringe SG, Tortell PD, Hallam SJ. 2009. Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. *Sci* 326:578–582. <https://doi.org/10.1126/science.1175309>
4. Canfield DE, Stewart FJ, Thamdrup B, De Brabandere L, Dalsgaard T, Delong EF, Revsbech NP, Ulloa O. 2010. A cryptic sulfur cycle in oxygen-minimum-zone waters off the Chilean coast. *Sci* 330:1375–1378. <https://doi.org/10.1126/science.1196889>
5. Zaikova E, Walsh DA, Stilwell CP, Mohn WW, Tortell PD, Hallam SJ. 2010. Microbial community dynamics in a seasonally anoxic fjord: Saanich Inlet, British Columbia. *Environ Microbiol* 12:172–191. <https://doi.org/10.1111/j.1462-2920.2009.02058.x>
6. Glaubitz S, Kießlich K, Meeske C, Labrenz M, Jürgens K. 2013. SUP05 dominates the gammaproteobacterial sulfur oxidizer assemblages in pelagic redoxclines of the central Baltic and Black seas. *Appl Environ Microbiol* 79:2767–2776. <https://doi.org/10.1128/AEM.03777-12>
7. Hawley AK, Brewer HM, Norbeck AD, Paša-Tolić L, Hallam SJ. 2014. Metaproteomics reveals differential modes of metabolic coupling among ubiquitous oxygen minimum zone microbes. *Proc Natl Acad Sci USA* 111:11395–11400. <https://doi.org/10.1073/pnas.1322132111>
8. Shah V, Morris RM. 2015. Genome sequence of “*Candidatus Thioglobus autotrophica*” strain EF1, a chemoautotroph from the SUP05 clade of marine gammaproteobacteria. *Genome Announc* 3:e01156-15. <https://doi.org/10.1128/genomeA.01156-15>
9. Shah V, Chang BX, Morris RM. 2017. Cultivation of a chemoautotroph from the SUP05 clade of marine bacteria that produces nitrite and consumes ammonium. *ISME J* 11:263–271. <https://doi.org/10.1038/ismej.2016.87>
10. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinform* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
11. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
12. Rodriguez-R LM, Konstantinidis KT. 2016. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *Peer J Preprints*. <https://doi.org/10.7287/peerj.preprints.1900v1>
13. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. *Nucleic Acids Res* 49:D1020–D1028. <https://doi.org/10.1093/nar/gkaa1105>