

Draft metagenome-assembled genome sequence of a *Dehalogenimonas* species from an enriched consortium for complete trichloroethylene reductive dechlorination

Lei Zhang,¹ Kaiwen Yang,¹ Xiaojun Zhang^{1,2}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT A draft metagenome-assembled genome was recovered from an anaerobic consortium capable of complete reductive dechlorination of trichloroethene to ethene. The draft genome, assigned to a *Dehalogenimonas* species, is 1.84 Mb in size with a G + C content of 54.53% and encodes 33 reductive dehalogenase homologs.

KEYWORDS dechlorination, consortium, metagenome, genome assembly

Organohalide-respiring bacteria conserve energy via reductive dehalogenases (RDase)-mediated respiratory reductive dehalogenation (1). However, only *Dehalococcoides* and *Dehalogenimonas* are known to completely dechlorinate trichloroethene (TCE) to ethene (1). Although *Dehalococcoides* genomes have been elucidated, genomic information on *Dehalogenimonas* remains limited. A *Dehalogenimonas* metagenome-assembled genome (MAG) was recovered from an anaerobic complete TCE dechlorination enrichment via dilution-to-extinction (2). The original enrichment was established from chloroethene-contaminated soil as previously described (3). The culture was serially diluted in DCB-1 mineral salt medium (2) amended with 20 mM acetate, 1.22 mM hydrogen, and 500 μ M TCE, and incubated at 30°C.

DNA was extracted using the DNeasy PowerSoil Pro Kit (QIAGEN, Cat. No. 47014). DNA quality was checked by Quantus fluorometer (PicoGreen). DNA was sheared to approximately 350 bp using a Covaris M220 system (Gene Company Limited, China). Paired-end libraries (NEXTFLEX Rapid DNA-Seq Kit) were sequenced on an Illumina NovaSeq X Plus platform (Illumina Inc., San Diego, CA, USA) using a NovaSeq X Series 25B Reagent Kit. The run generated 74.8 million paired-end reads (11.29 Gb). The raw reads were analyzed with anvi'o-8 platform (4). Reads quality control and filtering were performed using illumina-utils v2.13 (5). Unfiltered raw reads were deposited in the SRA; quality-filtered reads were used for downstream analyses. Taxonomic profiling used KrakenUniq v1.0.4 (6). Reads were assembled with MEGAHIT v1.2.9 (7) using default parameters, and contigs $\geq 1,000$ bp were retained. Reads were mapped with Bowtie2 v2.5.4 (8). Gene prediction used Prodigal v2.6.3 (9). Functional annotation was based on hidden Markov models (HMMs) against KEGG (10), Kofam (11), and Pfam (12) in anvi'o. Putative reductive dehalogenases were identified using the TIGR02486 HMM model. Binning used CONCOCT (13), MaxBin2 (14), and MetaBAT2 (15) with DASTool (16) refinement.

Dehalogenimonas MAG_dhgm_Z1 genome statistics are summarized in Table 1. Two contigs (totaling ~97 kb) identified as *Proteobacteria*-derived contaminants by NCBI's contamination screening were removed prior to final submission. A total of 1,858 coding sequences were predicted, including 33 putative RDase genes. These RDases were protein aligned with MAFFT (v7.526) (17) to reference RDases (18, 19) (Fig. 1A).

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

Address correspondence to Xiaojun Zhang,
xjzhang68@sjtu.edu.cn.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 13 November 2025

Accepted 19 February 2026

Published 12 March 2026

Copyright © 2026 Zhang 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/).

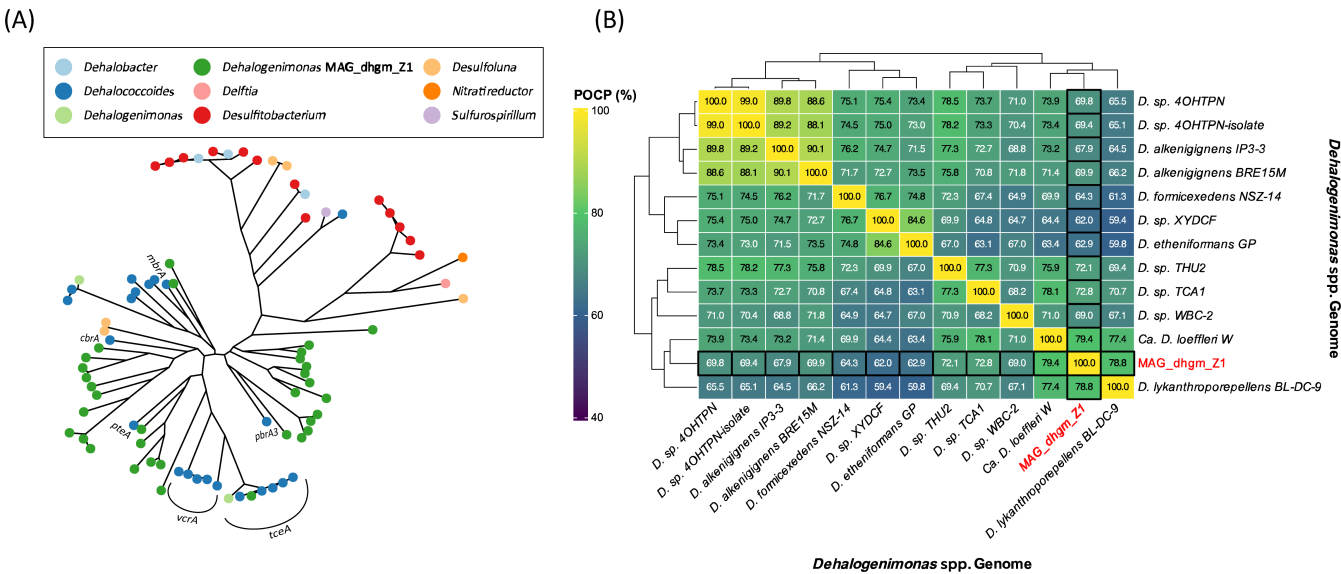


FIG 1 RDase phylogeny and Percentage of Conserved Proteins (POCP) relatedness of MAG_dhgm_Z1. (A) Maximum-likelihood phylogenetic tree of 33 putative RDases from MAG_dhgm_Z1 and reference RDases (18, 19) inferred with FastTree (v2.2.0) (20). (B) POCP comparisons among related genomes.

TABLE 1 Basic genomic features of *Dehalogenimonas* MAG_dhgm_Z1

No. of contigs	<i>N</i> ₅₀	GC content	Total size	Completeness	Redundancy
61	48,002	54.53%	1.84 Mb	97.18%	1.41%

HMM analysis identified one 16S rRNA gene and one 23S rRNA gene, and the 16S rRNA gene showed 99.46% identity with *Dehalogenimonas lykanthoporepellens* strain BL-DC-9^T (CP002084.1). Average nucleotide identity (ANI) and Percentage of Conserved Proteins (POCP) were calculated between the MAG and all available *Dehalogenimonas* complete genomes from NCBI using fastANI (21) v1.34 and POCP-nf pipeline (22, 23), respectively. The highest identity was 83.66% with *D. lykanthoporepellens* BL-DC-9^T, with lower similarities to other related strains, including Candidatus *D. loeffleri* strain W (78.04%), strain IP3-3^T (75.98%), and strain NSZ-14^T (75.6%). The POCP values supported assignment to the genus *Dehalogenimonas* (Fig. 1B). These values indicate that *Dehalogenimonas* MAG_dhgm_Z1 constitutes a distinct lineage within the genus *Dehalogenimonas*, expanding the *Dehalogenimonas* pangenome.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China project no. 32061133001.

AUTHOR AFFILIATIONS

¹State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China
²National Center for Translational Medicine (Shanghai), Shanghai Jiao Tong University, Shanghai, China

AUTHOR ORCIDs

Lei Zhang <http://orcid.org/0009-0001-2603-460X>
Xiaojun Zhang <http://orcid.org/0000-0002-4559-0177>

FUNDING

Funder	Grant(s)	Author(s)
National Natural Science Foundation of China	32061133001	Xiaojun Zhang

AUTHOR CONTRIBUTIONS

Lei Zhang, Formal analysis, Investigation, Methodology, Software, Writing – original draft
| Kaiwen Yang, Data curation, Software | Xiaojun Zhang, Conceptualization, Funding acquisition, Supervision, Writing – review and editing

DATA AVAILABILITY

The draft metagenome-assembled genome (MAG) sequence of *Dehalogenimonas* MAG_dhgm_Z1 has been deposited in European Nucleotide Archive (ENA) under accession [GCA_978046095.1](https://ena.ebi.ac.uk/ena/record/GCA_978046095.1). The associated ENA Study and Sample accessions are [PRJEB104910](https://ena.ebi.ac.uk/ena/record/PRJEB104910) and [SAMEA120968246](https://ena.ebi.ac.uk/ena/record/SAMEA120968246), respectively. The raw sequencing reads are available under ENA Run, accessions [ERR16125411](https://ena.ebi.ac.uk/ena/record/ERR16125411) and [ERR16125412](https://ena.ebi.ac.uk/ena/record/ERR16125412).

REFERENCES

- Hug LA, Maphosa F, Leys D, Löffler FE, Smidt H, Edwards EA, Adrian L. 2013. Overview of organohalide-respiring bacteria and a proposal for a classification system for reductive dehalogenases. *Phil Trans R Soc B* 368:20120322. <https://doi.org/10.1098/rstb.2012.0322>
- Löffler FE, Sanford RA, Ritalahti KM. 2005. Enrichment, cultivation, and detection of reductively dechlorinating bacteria. *Meth Enzymol* 397:77–111. [https://doi.org/10.1016/S0076-6879\(05\)97005-5](https://doi.org/10.1016/S0076-6879(05)97005-5)
- Yang K, Zhang L, Zhao K, Liu W, Tiehm A, Zhang X. 2026. Metabolism regulates spatial distribution patterns of different microbial taxonomic groups in chlorinated aliphatic hydrocarbons contaminated soil. *J Hazard Mater* 501:140640. <https://doi.org/10.1016/j.jhazmat.2025.140640>
- Eren AM, Kiehl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, et al. 2021. Community-led, integrated, reproducible multi-omics with anvio. *Nat Microbiol* 6:3–6. <https://doi.org/10.1038/s41564-020-00834-3>
- Eren AM, Vineis JH, Morrison HG, Sogin ML. 2013. A filtering method to generate high quality short reads using illumina paired-end technology. *PLoS One* 8:e66643. <https://doi.org/10.1371/journal.pone.0066643>
- Breitwieser FP, Baker DN, Salzberg SL. 2018. KrakenUniq: confident and fast metagenomics classification using unique k-mer counts. *Genome Biol* 19:198. <https://doi.org/10.1186/s13059-018-1568-0>
- Li D, Liu CM, Luo R, Sadakane K, Lam TW. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>
- Langmead B, Wilks C, Antonescu V, Charles R. 2019. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* 35:421–432. <https://doi.org/10.1093/bioinformatics/bty648>
- Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>
- Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30. <https://doi.org/10.1093/nar/28.1.27>
- Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. 2019. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 36:2251–2252. <https://doi.org/10.1101/602110>
- Paysan-Lafosse T, Andreeva A, Blum M, Chuguransky SR, Grego T, Pinto BL, Salazar GA, Bileschi ML, Llinares-López F, Meng-Papaxanthos L, Colwell LJ, Grishin NV, Schaeffer RD, Clementel D, Tosatto SCE, Sonnhammer E, Wood V, Bateman A. 2025. The Pfam protein families database: embracing AI/ML. *Nucleic Acids Res* 53:D523–D534. <https://doi.org/10.1093/nar/gkae997>
- Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. *Nat Methods* 11:1144–1146. <https://doi.org/10.1038/nmeth.3103>
- Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32:605–607. <https://doi.org/10.1093/bioinformatics/btv638>
- Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359. <https://doi.org/10.7717/peerj.7359>
- Siebert CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol* 3:836–843. <https://doi.org/10.1038/s41564-018-0171-1>
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>
- Molenda O, Puentes Jácome LA, Cao X, Nesbø CL, Tang S, Morson N, Patron J, Lomheim L, Wishart DS, Edwards EA. 2020. Insights into origins and function of the unexplored majority of the reductive dehalogenase gene family as a result of genome assembly and ortholog group classification. *Environ Sci: Processes Impacts* 22:663–678. <https://doi.org/10.1039/C9EM00605B>
- Zhou N, Li Q, Liang Z, Yu K, Zhang C, Wang H, Li P, He Z, Wang S. 2025. Microbially-mediated halogenation and dehalogenation cycling of organohalides in the ocean. *Nat Commun* 16:10670. <https://doi.org/10.1038/s41467-025-65696-x>
- 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>
- 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>
- Hölzer M. 2024. POCP-nf: an automatic Nextflow pipeline for calculating the percentage of conserved proteins in bacterial taxonomy. *Bioinformatics* 40:btac175. <https://doi.org/10.1093/bioinformatics/btac175>
- Qin Q-L, Xie B-B, Zhang X-Y, Chen X-L, Zhou B-C, Zhou J, Oren A, Zhang Y-Z. 2014. A proposed genus boundary for the prokaryotes based on genomic insights. *J Bacteriol* 196:2210–2215. <https://doi.org/10.1128/JB.01688-14>