



Complete Genome Sequence of *Azoarcus* sp. Strain DD4, a Gram-Negative Propanotroph That Degrades 1,4-Dioxane and 1,1-Dichloroethylene

Daiyong Deng,^a Fei Li,^a Lei Ye,^b  Mengyan Li^a

^aDepartment of Chemistry and Environmental Science, New Jersey Institute of Technology, Newark, New Jersey, USA

^bHaploX Biotechnology Co., Ltd., Shenzhen, Guangdong, China

ABSTRACT *Azoarcus* sp. strain DD4 can cometabolically degrade 1,4-dioxane and 1,1-dichloroethylene (1,1-DCE) when grown with propane and other substrates. The complete genome sequence of strain DD4 reveals a diverse collection of bacterial monooxygenase genes that may contribute to its versatility in degrading commingled groundwater pollutants.

Azoarcus sp. strain DD4 is a propanotrophic bacterium isolated from an activated sludge sample collected at a municipal wastewater treatment plant in northern New Jersey (1). Notably, DD4 presents a synchronic degradation ability for 1,4-dioxane and 1,1-dichloroethylene (1,1-DCE) via cometabolism with the induction of propane and some other substrates (1). Like other *Azoarcus* strains, DD4 is also a diazotroph that can assimilate atmospheric nitrogen (1–3). The growth and activity of DD4 can be sustained under a wide variety of aquifer-relevant conditions (1), suggesting that it has potential as an effective inoculum for *in situ* or *ex situ* bioaugmentation to treat the commingled contamination of 1,4-dioxane and 1,1-DCE. Therefore, the whole-genome sequence of DD4 provides insights into the genetic basis of its lifestyle and degradation capabilities, which are valuable to optimize and assess its field applications.

DD4 cells were harvested at the exponential phase after growth in nitrate mineral salts (NMS) medium with propane (0.1% [vol/vol] in the headspace) as the sole carbon and energy source. Total genomic DNA of DD4 was extracted using the MagAttract high-molecular-weight (HMW) DNA kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The extracted DNA was purified with AMPure PB magnetic beads and further used for the library preparation using the combination of the SMRTbell damage repair kit and barcoded adapter complete prep kit (Pacific Biosciences, Menlo Park, CA). The genome of DD4 was sequenced using the PacBio Sequel system, which generated approximately 1.69 Gbp of long-read sequencing data. The average length of raw sequences for sample DD4 is estimated as 2.1 kb, as the final number of raw reads is 802,558. Following the Hierarchical Genome Assembly Process (HGAP), a DD4 genome of high quality and accuracy was assembled using the RS_HGAP_Assembly.3 protocol and polished by Quiver in SMRT Portal v2.3.0 with default parameters (4). For genome component prediction, the GeneMarkS+ program (5) was employed to retrieve the related coding genes. Seven databases were then used for the annotation of gene functions (E value, $1E-5$; minimal alignment length percentage, greater than 40%) (6), namely, Gene Ontology (GO) (7), Kyoto Encyclopedia of Genes and Genomes (KEGG) (8), Clusters of Orthologous Groups (COG) (9), non-redundant protein (NR) databases (10), Transporter Classification Database (TCDB) (11), Swiss-Prot, and TrEMBL (12).

There exists one single circular chromosome in DD4 without circular or linear plasmids. The genome size of DD4 is 5,400,077 bp, with a GC content of 66.7%. A total

Citation Deng D, Li F, Ye L, Li M. 2019. Complete genome sequence of *Azoarcus* sp. strain DD4, a Gram-negative propanotroph that degrades 1,4-dioxane and 1,1-dichloroethylene. *Microbiol Resour Announc* 8:e00775-19. <https://doi.org/10.1128/MRA.00775-19>.

Editor David Rasko, University of Maryland School of Medicine

Copyright © 2019 Deng 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 Mengyan Li, Mengyan.Li@njit.edu.

Received 28 June 2019

Accepted 22 July 2019

Published 15 August 2019

of 5,001 putative genes are annotated, covering approximately 90.1% of the genome. The genome of DD4 contains 57 tRNA genes and 4 rRNA gene clusters (5S, 16S, and 23S). Five gene clusters encoding soluble di-iron monooxygenases (SDIMOs) (13, 14) are found on the chromosome. Based on phylogenetic analysis of the amino acid sequences of their hydroxylase alpha subunits, these five SDIMOs are categorized as two group 1 phenol hydroxylases, one group 2 toluene monooxygenase, one group 3 butane monooxygenase, and one group 5 propane monooxygenase. In addition, genes encoding a copper membrane particulate monooxygenase (15) and a cytochrome P450 CYP153 alkane hydroxylase are identified. One or more of these monooxygenases in DD4 may be responsible for initiating the oxidation of propane, 1,4-dioxane, 1,1-DCE, and other environmental pollutants (1, 16–19).

Data availability. The whole-genome sequence of *Azoarcus* sp. strain DD4 has been deposited in GenBank under the accession number [CP022958](https://doi.org/10.1093/nar/gkt1097). The raw reads have been deposited in SRA under the accession number [PRJNA398544](https://doi.org/10.1093/nar/gkt1097).

ACKNOWLEDGMENTS

This work was funded by the National Science Foundation (NSF; CAREER CBET-1846945), United States Geological Survey (USGS) State Water Resources Research Act Program (2018NJ400B), and the start-up fund from the Department of Chemistry and Environmental Science at the New Jersey Institute of Technology (NJIT). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

We declare no competing financial interests.

REFERENCES

- Deng D, Li F, Wu C, Li M. 2018. Synchronic biotransformation of 1,4-dioxane and 1,1-dichloroethylene by a Gram-negative propanotroph *Azoarcus* sp. DD4. *Environ Sci Technol Lett*. <https://doi.org/10.1021/acs.estlett.8b00312>.
- Krause A, Ramakumar A, Bartels D, Battistoni F, Bekel T, Boch J, Böhm M, Friedrich F, Hurek T, Krause L, Linke B, McHardy AC, Sarkar A, Schneider S, Syed AA, Thauer R, Vorhölter F-J, Weidner S, Pühler A, Reinhold-Hurek B, Kaiser O, Goesmann A. 2006. Complete genome of the mutualistic, N₂-fixing grass endophyte *Azoarcus* sp. strain BH72. *Nat Biotechnol* 24:1384. <https://doi.org/10.1038/nbt1243>.
- Faoro H, Rene Menegazzo R, Battistoni F, Gyaneshwar P, do Amaral FP, Taulé C, Rausch S, Gonçalves Galvão P, de Los Santos C, Mitra S, Heijo G, Sheu S-Y, Chen W-M, Mareque C, Zibetti Tadra-Sfeir M, Ivo Baldani J, Maluk M, Paula Guimarães A, Stacey G, de Souza EM, Pedrosa FO, Magalhães Cruz L, James EK. 2017. The oil-contaminated soil diazotroph *Azoarcus olearius* DQS-4^T is genetically and phenotypically similar to the model grass endophyte *Azoarcus* sp. BH72. *Environ Microbiol Rep* 9:223–238. <https://doi.org/10.1111/1758-2229.12502>.
- Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563. <https://doi.org/10.1038/nmeth.2474>.
- Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* 29:2607–2618. <https://doi.org/10.1093/nar/29.12.2607>.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene Ontology: tool for the unification of biology. *Nat Genet* 25:25. <https://doi.org/10.1038/75556>.
- Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40:D109–D114. <https://doi.org/10.1093/nar/gkr988>.
- Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4:41. <https://doi.org/10.1186/1471-2105-4-41>.
- Li W, Jaroszewski L, Godzik A. 2002. Tolerating some redundancy significantly speeds up clustering of large protein databases. *Bioinformatics* 18:77–82. <https://doi.org/10.1093/bioinformatics/18.1.77>.
- Saier MH, Jr, Reddy VS, Tamang DG, Västermark Å. 2014. The transporter classification database. *Nucleic Acids Res* 42:D251–D258. <https://doi.org/10.1093/nar/gkt1097>.
- Boeckmann B, Bairoch A, Apweiler R, Blatter M-C, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I. 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* 31:365–370. <https://doi.org/10.1093/nar/gkg095>.
- Leahy JG, Batchelor PJ, Morcomb SM. 2003. Evolution of the soluble diiron monooxygenases. *FEMS Microbiol Rev* 27:449–479. [https://doi.org/10.1016/S0168-6445\(03\)00023-8](https://doi.org/10.1016/S0168-6445(03)00023-8).
- Notomista E, Lahm A, Di Donato A, Tramontano A. 2003. Evolution of bacterial and archaeal multicomponent monooxygenases. *J Mol Evol* 56:435–445. <https://doi.org/10.1007/s00239-002-2414-1>.
- Coleman NV, Le NB, Ly MA, Ogawa HE, McCarl V, Wilson NL, Holmes AJ. 2012. Hydrocarbon monooxygenase in *Mycobacterium*: recombinant expression of a member of the ammonia monooxygenase superfamily. *ISME J* 6:171–182. <https://doi.org/10.1038/ismej.2011.98>.
- Deng D, Li F, Li M. 2018. A novel propane monooxygenase initiating degradation of 1,4-dioxane by *Mycobacterium dioxanotrophicus* PH-06. *Environ Sci Technol Lett* 5:86–91. <https://doi.org/10.1021/acs.estlett.7b00504>.
- Li M, Mathieu J, Yang Y, Fiorenza S, Deng Y, He Z, Zhou J, Alvarez PJ. 2013. Widespread distribution of soluble di-iron monooxygenase (SDIMO) genes in Arctic groundwater impacted by 1,4-dioxane. *Environ Sci Technol* 47:9950–9958. <https://doi.org/10.1021/es402228x>.
- Sales CM, Grostern A, Parales JV, Parales RE, Alvarez-Cohen L. 2013. Oxidation of the cyclic ethers 1,4-dioxane and tetrahydrofuran by a monooxygenase in two *Pseudonocardia* species. *Appl Environ Microbiol* 79:7702–7708. <https://doi.org/10.1128/AEM.02418-13>.
- Mahendra S, Alvarez-Cohen L. 2006. Kinetics of 1,4-dioxane biodegradation by monooxygenase-expressing bacteria. *Environ Sci Technol* 40:5435–5442. <https://doi.org/10.1021/es060714v>.