



Complete Genome Sequence of Halophilic Deep-Sea Bacterium *Halomonas axialensis* Strain Althf1

Megumi Tsurumaki,^{a,b} Senka Deno,^{a,c} Josephine Galipon,^{a,b}  Kazuharu Arakawa^{a,b,c}

^aInstitute for Advanced Biosciences, Keio University, Tsuruoka, Japan

^bSystems Biology Program, Graduate School of Media and Governance, Keio University, Fujisawa, Japan

^cFaculty of Environment and Information Studies, Keio University, Fujisawa, Japan

ABSTRACT *Halomonas axialensis* is a halophilic bacterial species discovered near a deep-sea hydrothermal vent. Here, we report the first single closed genome sequence of the original strain, *Halomonas axialensis* strain Althf1. The genome was assembled by Nanopore sequencing and consisted of a single chromosome of 3.6 Mbp with 56.8% G+C content.

Halomonas axialensis is a deep-sea halophilic bacterium most closely related to *Halomonas aquamarina* and *Halomonas meridian* (1). It was first isolated by Kaye et al. from low-temperature hydrothermal fluid (Cloud vent) at 1,530 m depth on the Axial Seamount on the Juan de Fuca Ridge in the Northeast Pacific Ocean (1, 2). Strain Althf1 was the first *H. axialensis* strain to be described. Also, a draft genome sequence containing 122 contigs (N_{50} value, 70,749 bp) for *H. axialensis* strain ACH-L-8, isolated from the South China Sea at 2,000 m depth, has been announced (3). Here, we report the first complete genome sequence (N_{50} value, 3,619,799 bp) of *H. axialensis* strain Althf1, obtained by Nanopore sequencing, to better understand deep-sea *Halomonas* biology.

H. axialensis strain Althf1 (ATCC BAA-802 [2], obtained directly from J. Z. Kaye) was cultured using ATCC medium (1097 *Halomonas* medium, with Casamino Acids replaced by high-molecular-weight casein). Total genomic DNA was extracted with a Genomic-tip 20/G (Qiagen), following the manufacturer's protocol. Sequencing libraries were prepared using the Rapid Barcoding Sequencing protocol (SQK-RBK004) provided by Oxford Nanopore Technologies. Long-read sequencing was performed using the GridION platform (Oxford Nanopore Technologies) with the FLO-MIN106 flow cell, and the quality of sequencing was monitored from the MinKNOW interface (Oxford Nanopore Technologies). Sequences were base called and demultiplexed using the Albacore software version 2.2.6 (Oxford Nanopore Technologies), and a total of 95,277 reads were obtained, with an average length of 5,471 bp. Reads with a length of 10,000 bp or more were retained for the assembly process, resulting in around 100× coverage. The resulting sequences were assembled with Canu version 1.7.1 (4), and the assembled single contig was manually circularized by eliminating an overlapping end. Assembly errors were corrected using nanopolish version 0.10.2. Assembly completeness was assessed by BUSCO v1 (5) on the gVolante server (6) and CheckM (7). Annotation was performed using DFAST v1.1.0 from the DDBJ (8). Two rounds of nanopolish error correction improved and saturated BUSCO completeness from 80% to 95% and CheckM completeness from 79.57% to 88.19%. The CheckM contamination level was 0.47%. The remaining indel rate was calculated by mapping the ACH-L-8 genes onto the Althf1 genome with minimap2 (9), and there were 2,959 deletions (0.11%) and 6,502 insertions (0.24%) out of the mapped 2,647,649-bp regions. Since the average length of a protein-coding sequence (CDS) is 933 bp, a CDS is likely to contain 1 deletion or 2

Citation Tsurumaki M, Deno S, Galipon J, Arakawa K. 2019. Complete genome sequence of halophilic deep-sea bacterium *Halomonas axialensis* strain Althf1. *Microbiol Resour Announc* 8:e00839-19. <https://doi.org/10.1128/MRA.00839-19>.

Editor J. Cameron Thrash, University of Southern California

Copyright © 2019 Tsurumaki 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 Kazuharu Arakawa, gaou@sfc.keio.ac.jp.

Received 12 July 2019

Accepted 12 July 2019

Published 1 August 2019

insertions, resulting in fragmented gene predictions. The remaining base mismatch error is difficult to assess, since there are variations within the strains, but it is estimated to be 2.40% at maximum (25,721 mismatches in 1,070,325 bp of 1,818 genes fully aligned between the two strains with BLASTN [10]). Note that shorter genes are more likely to be free of indels and are more likely to be aligned at full length. All software was run with default parameters.

The complete genome of this strain is 3,619,799 bp, with a G+C content of 56.8%. DFAST identified 5,503 predicted protein-coding sequences (CDS), 18 rRNA genes, and 62 tRNA genes. The genome of *H. axialensis* strain ACH-L-8 was previously reported to contain 20 aldehyde dehydrogenase (ALDH) genes (3) and assumed to be of interest for the bioremediation of aldehyde-contaminated environments. Our DFAST reanalysis of strain ACH-L-8 confirmed the presence of 17 such ALDH coding sequences, 16 of which were also detected by BLASTN 2.8.1+ (10) in the newly sequenced *H. axialensis* strain Althf1. Both strains harbored genes involved in resistance to mercury and arsenic.

This is the first report of a complete genome sequence of *H. axialensis*, as well as the first report for the strain that initially enabled the identification of this species (2). This assembly will facilitate genome-wide comparison studies with a focus on the ecology of bacteria living in deep-sea hydrothermal environments.

Data availability. The chromosome and plasmid sequences reported here were deposited in the DDBJ under the accession number [AP019517](#) and in the Sequence Read Archive (SRA) under BioProject accession number [PRJNA521630](#).

ACKNOWLEDGMENTS

We thank J. Z. Kaye, the original discoverer of the strain, for donating one of his old stocks, as well as Yuki Yoshida and Yuki Takai for technical assistance.

Sequencing and assembly were conducted during the Genome Engineering Workshop course of the Systems Biology Program in the Graduate School of Media and Governance at Keio University.

This work was supported, in part, by research funds from the Yamagata Prefectural Government and Tsuruoka City, Japan.

REFERENCES

1. Kaye JZ, Baross JA. 2000. High incidence of halotolerant bacteria in Pacific hydrothermal-vent and pelagic environments. *FEMS Microbiol Ecol* 32:249–260. <https://doi.org/10.1111/j.1574-6941.2000.tb00718.x>.
2. Kaye JZ, Márquez MC, Ventosa A, Baross JA. 2004. *Halomonas neptunia* sp. nov., *Halomonas sulfidaeris* sp. nov., *Halomonas axialensis* sp. nov. and *Halomonas hydrothermalis* sp. nov.: Halophilic bacteria isolated from deep-sea hydrothermal-vent environments. *Int J Syst Evol Microbiol* 54:499–511. <https://doi.org/10.1099/ijs.0.02799-0>.
3. Ye J, Ren C, Shan X, Zeng R, Pro D. 2016. Draft genome sequence of aldehyde-degrading strain *Halomonas axialensis* ACH-L-8. *Genome Announc* 4:e00287–16. <https://doi.org/10.1128/genomeA.00287-16>.
4. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2016. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
5. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.
6. Nishimura O, Hara Y, Kuraku S. 2017. gVolante for standardizing completeness assessment of genome and transcriptome assemblies. *Bioinformatics* 33:3635–3637. <https://doi.org/10.1093/bioinformatics/btx445>.
7. 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>.
8. Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34:1037–1039. <https://doi.org/10.1093/bioinformatics/btx713>.
9. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
10. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.