AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Halomonas sulfidaeris* Strain Esulfide1 Isolated from a Metal Sulfide Rock at a Depth of 2,200 Meters, Obtained Using Nanopore Sequencing

Motofumi Saito,[a,b] Akane Nishigata,[a,b] Josephine Galipon,[a,b] Kazuharu Arakawa[a,b,c]

[a]Institute for Advanced Biosciences, Keio University, Tsuruoka, Japan
[b]Systems Biology Program, Graduate School of Media and Governance, Keio University, Fujisawa, Japan
[c]Faculty of Environment and Information Studies, Keio University, Fujisawa, Japan

ABSTRACT We report the complete genome sequence of *Halomonas sulfidaeris* ATCC BAA-803, isolated from a metal sulfide rock at a depth of 2,200 m in the Northeast Pacific Ocean. The assembled genome comprised one circular chromosome of 4.20 Mb and one large plasmid of 273 kb. The chromosome harbors 6,705 protein-coding genes.

Halomonas sulfidaeris strain ATCC BAA-803 (=CECT 5817 = DSM 15722) is a Gram-negative rod-shaped aerobic bacterium classified in the phylum *Proteobacteria*, order *Oceanospirillales*, and family *Halomonadaceae*. It was first isolated from metal sulfide rock at a 2,200-m depth in the Main Endeavor Field of the Endeavor Segment of the Juan de Fuca Ridge, Northeast Pacific Ocean, as strain Esulfide1 (1). It grows optimally at 30°C (pH 7) with 18% total salts. Deep-sea halomonids are of interest for their ability to grow at a wide range of temperatures and potential for seafloor bioremediation. However, very few complete genomes are yet available for these organisms. *H. sulfidaeris* Esulfide1 is of particular interest due to its proximity to hydrothermal vents and the type of rock from which it was sampled.

The strain was cultured using ATCC medium (1097 *Halomonas* medium, with Casamino Acids replaced by high-molecular-weight casein). The liquid culture was passaged two times overnight at 30°C in a shaker and harvested at an optical density at 600 nm (OD_{600}) of ≈0.6. Genomic DNA was extracted using a Genomic-tip 20/G kit (Qiagen). A sequencing library was prepared with a rapid barcoding kit (SQK-RBK004) and was subsequently sequenced using an R9.4 flow cell (FLO-MIN 106) on a GridION X5 system (Oxford Nanopore Technologies).

Adapters were removed, and all 98,820 reads were demultiplexed using Porechop version 0.2.3. The average read length was 4,484 bp, with an N_{50} value of 8,468 bp (maximum read length, 109,205 bp) for a coverage of 98×. The reads were assembled using Canu version 1.7.1 (2), and the overlapping end was manually removed to finish circularization. The reads were mapped to the draft 4.20-Mb Nanopore genome using the Burrows-Wheeler Aligner "mem" (BWA-mem) algorithm version 0.7.17 (3). The alignment file was then sorted and indexed with SAMtools (4), and consensus was called using Racon (5). The genome sequence was further polished with Nanopolish version 0.9.2 (6) with minimap v2.5-r622-dirty (7). The quality of the assembly was assessed by calculating the genome completeness using gVolante version 1.2.0 (8), which showed 87.5% BUSCO completeness, indicating a certain amount of uncorrected indels. The genome sequence was annotated by DFAST (9). All software programs were used with default settings.

H. sulfidaeris strain Esulfide1 (ATCC BAA-803) has a circular chromosome of 4,207,221 bp and one plasmid with a length of 273,549 bp. The chromosome has a G+C

Citation Saito M, Nishigata A, Galipon J, Arakawa K. 2019. Complete genome sequence of *Halomonas sulfidaeris* strain Esulfide1 isolated from a metal sulfide rock at a depth of 2,200 meters, obtained using Nanopore sequencing. Microbiol Resour Announc 8:e00327-19. https://doi.org/10.1128/MRA.00327-19.

Editor David Rasko, University of Maryland School of Medicine

Address correspondence to Kazuharu Arakawa, gaou@sfc.keio.ac.jp.

Received 19 March 2019
Accepted 9 May 2019
Published 6 June 2019

content of 53.8% and contains 6,785 predicted genes encoding 6,705 proteins, 61 tRNAs, and 18 rRNAs. The genome size was similar to that of the contig-level assembly of *H. sulfidaeris* strain SST4, isolated from a trench in South Shetland (NCBI RefSeq accession no. NZ_QNTU00000000). Strain ATCC BAA-803 additionally contained one plasmid that is absent from SST4, with a length of 273 kb and a G+C content of 53.4%. The G+C contents of the chromosome and plasmid were equivalent, and the plasmid shared homology with the plasmid of *Halomonas* sp. strain KO116 (GenBank acccession no. CP011053) and the plasmid of *Halomonas ventosae* strain NRS2HaP1 (GenBank accession no. CP022738), albeit 40 kb shorter and 155 kb longer, respectively. It contained 412 protein-coding genes, 400 of which were of unknown function (hypothetical protein). Interestingly, out of the 12 coding sequences for which a function could be predicted, 4 were involved in resistance to heavy metals, including mercury, cadmium, and lead. The new plasmid also harbored an IS*256* family transposase that was absent from both CP011053 and CP022738, an indication of active transposable elements (10). The completeness and quality of this genome will allow genome-wide comparison analysis with other species of *Halomonas*.

Data availability. The chromosome and plasmid sequences reported here were deposited in DDBJ/ENA under the accession no. AP019514 and AP019515 and in the Sequence Read Archive (SRA) under BioProject accession no. PRJNA521444.

ACKNOWLEDGMENTS

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

The sequencing and assembly were conducted in the Genome Engineering Workshop course of the Systems Biology Program, Graduate School of Media and Governance, Keio University.

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

REFERENCES

1. 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.
2. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. 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.
3. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
4. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
5. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116.
6. Loman NJ, Quick J, Simpson JT. 2015. A complete bacterial genome assembled de novo using only Nanopore sequencing data. Nat Methods 12:733. https://doi.org/10.1038/nmeth.3444.
7. Li H. 2017. Minimap2: pairwise alignment for nucleotide sequences. arXiv arXiv:1708.01492 [q-bio.GN]. https://arxiv.org/abs/1708.01492.
8. 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.
9. Tanizawa Y, Fujisawa T, Kaminuma E, Nakamura Y, Arita M. 2016. DFAST and DAGA: Web-based integrated genome annotation tools and resources. Food Health 35:173–184. https://doi.org/10.12938/bmfh.16-003.
10. Hennig S, Ziebuhr W. 2010. Characterization of the transposase encoded by IS256, the prototype of a major family of bacterial insertion sequence elements. J Bacteriol 192:4153–4163. https://doi.org/10.1128/JB.00226-10.