





Complete Genome Sequence of *Halomonas olivaria*, a Moderately Halophilic Bacterium Isolated from Olive Processing Effluents, Obtained by Nanopore Sequencing

 Shohei Nagata,^{a,b}  Konosuke M. Ii,^{a,b}  Tomoya Tsukimi,^{a,b}  Masahiro C. Miura,^{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 Here, we report the complete genome sequence of *Halomonas olivaria* strain TYRC17, a moderately halophilic, Gram-negative bacterium that was isolated from olive processing effluents. The 5-Mbp genome consists of 7,375 protein-coding sequences, including a variety of genes involved in tyrosol metabolism, nitrate respiration, and the production of polysaccharides.

Halomonas is a genus of Gram-negative halophilic bacteria widely distributed in high-salinity environments, and its members typically require a high NaCl concentration and alkaline pH for growth (1). Many species within the genus are known to have potential applications in biotechnology due to their capacity to produce osmoprotectants, as well as certain enzymes or exopolysaccharides, and for their role in the degradation of pollutants (1–3). A strain isolated from salted olive processing effluents, *Halomonas olivaria*, was shown to completely degrade tyrosol, a toxic compound found in such effluents (4, 5).

H. olivaria strain TYRC17, from the Mediterranean Institute of Oceanography (Marseille, France) and also available from the Leibniz Institute DSMZ–German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany), was grown as described previously (4). When the optical density (OD) reached 0.3, the culture solution was centrifuged at $5,000 \times g$ for 10 min at 20°C to pellet the cells. Genomic DNA was extracted and purified with the Genomic-tip 20/G kit following the manufacturer's instructions (Qiagen, Hilden, Germany). The sequencing libraries were prepared using a rapid barcoding kit (SQK-RAB004) and sequenced in an R9.4 flow cell (FLO-MIN106) with the GridION device (Oxford Nanopore Technologies).

A total of 1,000,874 reads were obtained, with lengths ranging from 5 to 284,063 bp. After filtering the reads to retain those greater than 50,000 bp to bring coverage to just below $100\times$, filtered reads were assembled using Canu version 1.8 (6), using the default parameters for Nanopore data. The genome was assembled into a single contig and was manually circularized by deleting an overlapping end. An improved consensus sequence for the assembly was obtained with Nanopolish software version 0.10.2 (7), using the methylation-aware option. Genome completeness was estimated using Benchmarking Universal Single-Copy Orthologs (BUSCO) software version 1 (8) with the bacterial data set on the gVolante Web server (9). Of the 40 BUSCO group genes, 95% were completely (85%) or partially (10%) recovered in the polished assembly, indicating a certain amount of uncorrected indels. The comparison of six 16S rRNA genes in our genome with the known partial 16S rRNA sequence (1,504 bp) indicated $99.22 \pm 0.12\%$ identity. Genes were annotated using the DDBJ Fast Annotation and Submission Tool (DFAST) pipeline (10). The estimated genome size is 4,998,083 bp, with a G+C content

Citation Nagata S, Ii KM, Tsukimi T, Miura MC, Galipon J, Arakawa K. 2019. Complete genome sequence of *Halomonas olivaria*, a moderately halophilic bacterium isolated from olive processing effluents, obtained by Nanopore sequencing. Microbiol Resour Announc 8:e00144-19. <https://doi.org/10.1128/MRA.00144-19>.

Editor Vincent Bruno, University of Maryland School of Medicine

Copyright © 2019 Nagata 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.

S.N., K.M.I., T.T., and M.C.M. contributed equally to this work.

Received 11 February 2019

Accepted 15 April 2019

Published 2 May 2019

of 55.3%. The genome was predicted to contain 7,375 putative coding sequences, 18 rRNAs, 59 tRNAs, and 1 CRISPR sequence.

In our analysis, four nitroreductases (Cluster of Orthologous Groups 0778 [COG0778]), one nitrous oxidase accessory protein (COG3420), and one P_{II} nitrogen regulatory protein (COG0347) were found. In addition, concerning sugar metabolism, 40 genes corresponding to ABC-type sugar transport systems (COG1129, COG1175, COG1879, and COG3839) were annotated. Finally, among the tyrosol metabolic pathways shown in reference 4, a gene coding for homogentisate 1,2-dioxygenase and partial sequences associated with aryl-alcohol dehydrogenation or 4-hydroxyphenylacetic acid degradation were detected. This information will be useful to reveal the mechanism of tyrosol degradation mediated by *H. olivaria* and to contribute to the field of bioremediation.

Data availability. The complete genome sequence of *H. olivaria* has been deposited in DDBJ under the accession number [AP019416](https://www.ncbi.nlm.nih.gov/nuccore/AP019416) and in the Sequence Read Archive (SRA) under the accession number [PRJNA521098](https://www.ncbi.nlm.nih.gov/sra/PRJNA521098).

ACKNOWLEDGMENTS

We thank the Mediterranean Institute of Oceanography (Marseille, France) for kindly donating the strain and 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 from Tsuruoka City, Japan.

REFERENCES

- García MT, Ventosa A, Mellado E. 2005. Catabolic versatility of aromatic compound-degrading halophilic bacteria. *FEMS Microbiol Ecol* 54: 97–109. <https://doi.org/10.1016/j.femsec.2005.03.009>.
- Piubeli F, Grossman MJ, Fantinatti-Garboggini F, Durrant LR. 2012. Enhanced reduction of COD and aromatics in petroleum-produced water using indigenous microorganisms and nutrient addition. *Int Biodeterior Biodegrad* 68:78–84. <https://doi.org/10.1016/j.ibiod.2011.11.012>.
- Zhuang X, Han Z, Bai Z, Zhuang G, Shim H. 2010. Progress in decontamination by halophilic microorganisms in saline wastewater and soil. *Environ Pollut* 158:1119–1126. <https://doi.org/10.1016/j.envpol.2010.01.007>.
- Liebgott PP, Labat M, Amouric A, Tholozan JL, Lorquin J. 2008. Tyrosol degradation via the homogentisic acid pathway in a newly isolated *Halomonas* strain from olive processing effluents. *J Appl Microbiol* 105: 2084–2095. <https://doi.org/10.1111/j.1365-2672.2008.03925.x>.
- Amouric A, Liebgott PP, Joseph M, Brochier-Armanet C, Lorquin J. 2014. *Halomonas olivaria* sp. nov., a moderately halophilic bacterium isolated from olive-processing effluents. *Int J Syst Evol Microbiol* 64:46–54. <https://doi.org/10.1099/ijs.0.049007-0>.
- 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>.
- Loman NJ, Quick J, Simpson JT. 2015. A complete bacterial genome assembled *de novo* using only Nanopore sequencing data. *Nat Methods* 12:733–735. <https://doi.org/10.1038/nmeth.3444>.
- 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>.
- 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>.
- 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>.