Complete Genome Sequence of *Nitrosomonas cryotolerans* ATCC 49181, a Phylogenetically Distinct Ammonia-Oxidizing Bacterium Isolated from Arctic Waters

Marlen C. Rice, Jeanette M. Norton, Lisa Y. Stein, Jessica Kozlowski, Annette Bollmann, Martin G. Klotz, Luis Sayavedra-Soto, Nicole Shapiro, Lynne A. Goodwin, Marcel Huntemann, Alicia Clum, Manoj Pillay, Neha Varghese, Natalia Mikhailova, Krishna Palaniappan, Natalia Ivanova, Supratim Mukherjee, T. B. K. Reddy, Chew Yee Ngan, Chris Daum, Nikos Kyrpides, Tanja Woyke

Utah State University, Logan, Utah, USA; University of Alberta, Edmonton, Alberta, Canada; Miami University, Oxford, Ohio, USA; Queens College in The City University of New York, Flushing, New York, USA; The Institute of Marine Microbes and Ecospheres, Xiamen University, Xiamen, China; Oregon State University, Corvallis, Oregon, USA; DOE Joint Genome Institute, Walnut Creek, California, USA; Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico, USA.

**ABSTRACT** *Nitrosomonas cryotolerans* ATCC 49181 is a cold-tolerant marine ammonia-oxidizing bacterium isolated from seawater collected in the Gulf of Alaska. The high-quality complete genome contains a 2.87-Mbp chromosome and a 56.6-kbp plasmid. Chemolithoautotrophic modules encoding ammonia oxidation and CO₂ fixation were identified.

*Nitrosomonas cryotolerans* ATCC 49181 was isolated from surface seawater at 59°47'67"N, 151°55'08"W in 1980 and described in 1988 (1). *Nitrosomonas cryotolerans* is an obligate halophilic betaproteobacterium in the *Nitrosomonadaceae* (2) that is able to grow at temperatures as low as -5°C (3), the coldest temperature known for any bacterium in the *Nitrosomonadaceae*. *Nitrosomonas cryotolerans* was selected for sequencing based on phylogenetic and physiologic interests. Genome comparisons with *N. cryotolerans* will lead to better delineation of the *Nitrosomonas* genus.

*Nitrosomonas cryotolerans* was grown in marine medium containing 25 mM ammonium at 15°C (4). Cultures were harvested by centrifugation, and genomic DNA (gDNA) was isolated using the JGI cetyltrimethylammonium bromide (CTAB) protocol. The quality of gDNA was checked by gel electrophoresis and by amplification and sequencing of the V4 region of the small subunit (SSU) rRNA gene (5).

Genomic DNA was sequenced at the Department of Energy (DOE) Joint Genome Institute (JGI) using Pacific Biosciences (PacBio) technology (6). A PacBio SMRTbell library was constructed and sequenced on the PacBio RS platform, which generated 178,102 reads totaling 560.1 Mb. The input read coverage was 191.1 X. Raw reads were assembled using HGAP version 2.2.0.p1 (7). The final assembly contains two contigs (one chromosome and one plasmid) totaling 2.87 Mbp in size. An earlier version accomplished using the Illumina HiSeq 2000 platform remained in 91 scaffolds, with nearly identical repeats of key gene clusters unresolved, and was not combined with the PacBio sequence for the complete genome reported here.

Genes were identified using Prodigal (8), using the DOE-JGI microbial annotation pipeline (9), followed by manual curation using GenePRIMP (10). The predicted coding sequences (CDSs) were translated and used to search the NCBI nonredundant, UniProt,

Received 10 January 2017  Accepted 12 January 2017  Published 16 March 2017

Citation Rice MC, Norton JM, Stein LY, Kozlowski J, Bollmann A, Klotz MG, Sayavedra-Soto L, Shapiro N, Goodwin LA, Huntemann M, Clum A, Pillay M, Varghese N, Mikhailova N, Palaniappan K, Ivanova N, Mukherjee S, Reddy TBK, Yee Ngan C, Daum C, Kyrpides N, Woyke T. 2017. Complete genome sequence of *Nitrosomonas cryotolerans* ATCC 49181, a phylogenetically distinct ammonia-oxidizing bacterium isolated from Arctic waters. Genome Announc 5:e00011-17. https://doi.org/10.1128/genomeA.00011-17.

Copyright © 2017 Rice et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Jeanette M. Norton, jeannette.norton@usu.edu.
The genome of *N. cryotolerans* contains one ribosomal operon (SAMN02743940_1059 to SAMN02743940_1062). Metabolic modules encoding chemolithotrophic ammonia metabolism include three clusters each of ammonia-monoxygenase and hydroxylamine-oxidoreductase genes. The *amo* gene clusters have unique arrangements of neighboring genes, including the *amoCABDcopCD* cluster (SAMN02743940_0955 to SAMN02743940_0961), *amoCABED* cluster (SAMN02743940_1533 to SAMN02743940_1537), and *amoCA* cluster (SAMN02743940_2438 to SAMN02743940_2436). The genome contains 2 standalone *amoC* genes (SAMN02743940_2312 and 2350). Three *haoABcycAB* clusters are located at SAMN02743940_0346 to SAMN02743940_0349, SAMN02743940_0686 to SAMN02743940_0689, and SAMN02743940_1062. The *cryotolerans* genome contains a single gene cluster encoding the Calvin-Benson-Bassham cycle for carbon assimilation (SAMN02743940_1712 to SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO) with highest nucleic acid sequence identity to homologous genes in *N. cryotolerans* (SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO) with highest nucleic acid sequence identity to homologous genes in *N. cryotolerans* (SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO) with highest nucleic acid sequence identity to homologous genes in *N. cryotolerans* (SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO). The genome contains 2 standalone *amoC* genes (SAMN02743940_2312 and 2350). Three *haoABcycAB* clusters are located at SAMN02743940_0346 to SAMN02743940_0349, SAMN02743940_0686 to SAMN02743940_0689, and SAMN02743940_1062. The *cryotolerans* genome contains a single gene cluster encoding the Calvin-Benson-Bassham cycle for carbon assimilation (SAMN02743940_1712 to SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO) with highest nucleic acid sequence identity to homologous genes in *N. cryotolerans* (SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO). The genome contains 2 standalone *amoC* genes (SAMN02743940_2312 and 2350). Three *haoABcycAB* clusters are located at SAMN02743940_0346 to SAMN02743940_0349, SAMN02743940_0686 to SAMN02743940_0689, and SAMN02743940_1062. The *cryotolerans* genome contains a single gene cluster encoding the Calvin-Benson-Bassham cycle for carbon assimilation (SAMN02743940_1712 to SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO) with highest nucleic acid sequence identity to homologous genes in *N. cryotolerans* (SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO). The genome contains 2 standalone *amoC* genes (SAMN02743940_2312 and 2350). Three *haoABcycAB* clusters are located at SAMN02743940_0346 to SAMN02743940_0349, SAMN02743940_0686 to SAMN02743940_0689, and SAMN02743940_1062. The *cryotolerans* genome contains a single gene cluster encoding the Calvin-Benson-Bassham cycle for carbon assimilation (SAMN02743940_1712 to SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO) with highest nucleic acid sequence identity to homologous genes in *N. cryotolerans* (SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO). The genome contains 2 standalone *amoC* genes (SAMN02743940_2312 and 2350). Three *haoABcycAB* clusters are located at SAMN02743940_0346 to SAMN02743940_0349, SAMN02743940_0686 to SAMN02743940_0689, and SAMN02743940_1062. The *cryotolerans* genome contains a single gene cluster encoding the Calvin-Benson-Bassham cycle for carbon assimilation (SAMN02743940_1712 to SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO) with highest nucleic acid sequence identity to homologous genes in *N. cryotolerans* (SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO). The genome contains 2 standalone *amoC* genes (SAMN02743940_2312 and 2350). Three *haoABcycAB* clusters are located at SAMN02743940_0346 to SAMN02743940_0349, SAMN02743940_0686 to SAMN02743940_0689, and SAMN02743940_1062. The *cryotolerans* genome contains a single gene cluster encoding the Calvin-Benson-Bassham cycle for carbon assimilation (SAMN02743940_1712 to SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO) with highest nucleic acid sequence identity to homologous genes in *N. cryotolerans* (SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO). The genome contains 2 standalone *amoC* genes (SAMN02743940_2312 and 2350). Three *haoABcycAB* clusters are located at SAMN02743940_0346 to SAMN02743940_0349, SAMN02743940_0686 to SAMN02743940_0689, and SAMN02743940_1062. The *cryotolerans* genome contains a single gene cluster encoding the Calvin-Benson-Bassham cycle for carbon assimilation (SAMN02743940_1712 to SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO) with highest nucleic acid sequence identity to homologous genes in *N. cryotolerans* (SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO). The genome contains 2 standalone *amoC* genes (SAMN02743940_2312 and 2350). Three *haoABcycAB* clusters are located at SAMN02743940_0346 to SAMN02743940_0349, SAMN02743940_0686 to SAMN02743940_0689, and SAMN02743940_1062. The *cryotolerans* genome contains a single gene cluster encoding the Calvin-Benson-Bassham cycle for carbon assimilation (SAMN02743940_1712 to SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO) with highest nucleic acid sequence identity to homologous genes in *N. cryotolerans* (SAMN02743940_1708 LysR, CbbL, CbbS, CbbQ, and CbbO). The genome contains 2 standalone *amoC* genes (SAMN02743940_2312 and 2350).
identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.

9. Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Pala

niappan K, Szeto E, Pillay M, Chen I-M, Pati A, Nielsen T, Markowitz VM, Kyr

pides NC. 2015. The standard operating procedure of the DOE-JGI microbial genome annotation pipeline (MGAP v. 4). Stand Genomic Sci 10:86.

10. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. 2010. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. Nat Methods 7:455–457. https://doi.org/10.1038/nmeth.1457.

11. Lowe TM, Eddy SR. 1997. TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964. https://doi.org/10.1093/nar/25.5.0955.

12. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Göckner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196. https://doi.org/10.1093/nar/gkm864.

13. Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics 29:2933–2935. https://doi.org/10.1093/bioinformatics/btt509.

14. Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Pillay M, Ratner A, Huang J, Woyke T, Huntemann M. 2013. IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res 42:D560–D567.

15. Chen IM, Markowitz VM, Palaniappan K, Szeto E, Chu K, Huang J, Ratner A, Pillay M, Hadjithomas M, Huntemann M, Mikhailova N, Ovchinnikova G, Ivanova NN, Kyrpides NC. 2016. Supporting community annotation and user collaboration in the integrated microbial genomes (IMG) system. BMC Genomics 17:307. https://doi.org/10.1186/s12864-016-2629-y.

16. Kozłowski JA, Kits KD, Stein LY. 2016. Complete genome sequence of Nitrosomonas ureae strain Nm10, an oligotrophic group 6a nitrosonomad. Genome Announc 4(2):e00094-16. https://doi.org/10.1128/genomeA.00094-16.

17. Suva Y, Norton JM, Bollmann A, Klotz MG, Stein LY, Laanbroek HJ, Arp DJ, Goodwin LA, Chertkov O, Held B, Bruce D, Detter JC, Detter JC, Tapia R, Han CS. 2011. Genome sequence of Nitrosomonas sp. strain AL212, an ammonia-oxidizing bacterium sensitive to high levels of ammonia. J Bacteriol 193:5047–5048. https://doi.org/10.1128/JB.05521-11.