Draft genome sequence of *Nitrosomonas sp.* ANs5, an extremely alkalitolerant ammonia-oxidizing bacterium isolated from Mongolian soda lakes

Jose R. Valera[a], Dimitry Y. Sorokin [b,c], Alyson E. Santoro[a#]

[a]University of California, Department of Ecology, Evolution, & Marine Biology, Santa Barbara, California, USA

[b]Winogradsky Institute of Microbiology, Research Centre of Biotechnology, Russian Academy of Sciences, Moscow, Russia

[c]Department of Biotechnology, TU Delft, The Netherlands

Running head: *Genome of Nitrosomonas sp.* ANs5

[#]Address correspondence to: Alyson Santoro, asantoro@ucsb.edu

Abstract

The draft genome of a chemolithoautotrophic ammonia-oxidizing bacterium of the genus Nitrosomonas *is reported.* Nitrosomonas sp. *strain ANs5, previously classified as a strain of* N. halophila*, is an alkalitolerant ammonia-oxidizing bacterium isolated from the soda lakes of northeast Mongolia.*

Announcement

Nitrosomonas sp. strain ANs5 (hereafter ANs5) is one of five nitrosomonads isolated from composite sediment samples of the saline soda lakes in the northeast region of Mongolia (Choibolsan Province) (1). ANs5 is a halotolerant, ammonia-oxidizing bacterium (AOB) in the order of *Burkholderiales*, in the family of *Nitrosomonadaceae* (Betaproteobacteria) that is capable of growth at pH values as high as 11.4 (1), the highest pH known for any bacterium in the *Nitrosomonadaceae*. ANs5 was originally classified as a strain of *Nitrosomonas halophila*, of which the type strain is strain Nm1, and expanded the species description to include alkali-tolerant properties (2). Genome comparisons between ANs5, Nm1 and non-alkaliphilic strains may lead to better understanding of alkaliphilic adaptations within the AOB.

ANs5 was grown in batch culture on a temperature-controlled shaker set to 100 rpm, at 30˚C in 160 mL glass screw top bottles containing 30 mL of media. The alkaline (pH = 9.7-10) medium was prepared as previously described, with 10 mM ammonium (1). Eight cultures were combined and collected via vacuum filtration on a 25 mm, 0.22 µm pore sized polyethersulfone (Pall Supor) membrane filter. DNA was extracted using a modification of the DNeasy Blood & Tissue kit (Qiagen) as previously described (3). DNA libraries were constructed using the Illumina DNA Prep kit and sequenced using the Illumina NovaSeq X Plus platform with 2 x 151 bp paired-end reads, producing a total of 1,541,784 paired reads.

The draft genome was analyzed using the Kbase open-source platform for genome assembly and analysis (4). Default parameters were used except where otherwise noted. Quality control and pre-processing were conducted with FASTQC (v 0.12.1) (5), TRIMMomatic (v 0.36; sliding window = 5, min. quality = 20) (6), and PRINSeq (v 0.20.4) (7), in that order. Processed reads were assembled using Unicycler (v 0.4.8; min. contig length = 2000 bp, contig bridging threshold = 'bold') (8). The assembled genome (read coverage = 65x) was annotated via NCBI PGAP (v 6.8) (9). Our phylogenetic analysis, based on a concatenated alignment of single-copy genes in the Proteobacteria HMM set (n = 119 genes) within GToTree (v 1.8.10) (10) was used to construct the maximum-likelihood tree with FastTree2 (v 2.1.11) (11), visualized with iToL (v 7.2) (12) (Figure 1).

The final assembly contains 122 contigs with an N_{50} value of 49,119, totaling 3.07 Mbp, and a GC content of 52% (Table 1). The genome contained 2,785 protein-coding genes and is 99.82% complete with 0.42% contamination, estimated by CheckM2 (v 1.1.0) (13). Genes encoding chemolithotrophic ammonia oxidation were identified, including two copies of the ammonia monooxygenase subunits (*amoCAB*; ABTW62_11265:75 & ABTW62_07675:85) and a single copy of the hydroxylamine oxidoreductase (*hao*; ABTW62_13545) gene. We identified the gene encoding nitrite reductase (*nirK*; ABTW62_02095). Strains ANs5 and Nm1 share an average nucleotide identity (ANI) of 91.9% as determined by FASTANI (v 0.1.3) (14). Using the Compare Two Proteomes tool in KBase with a <60% amino acid identity threshold for unique genes, we identified 500 genes unique to ANs5 that are not shared with Nm1.

Table 1. Genome features of *Nitrosomonas* sp. strain ANs5

Strain	NCBI Accession Number	Genome size (bp)	GC (%)	No. of contigs	No. of total genes	No. of protein CDS	No. of rRNA operons	No. of tRNA operons	Completeness / Contamination (%)
Nitrosomonas sp. ANs5	PRJNA1125436	3,066,906	52.34	122	3,395	3,363	2	38	99.82 / 0.42

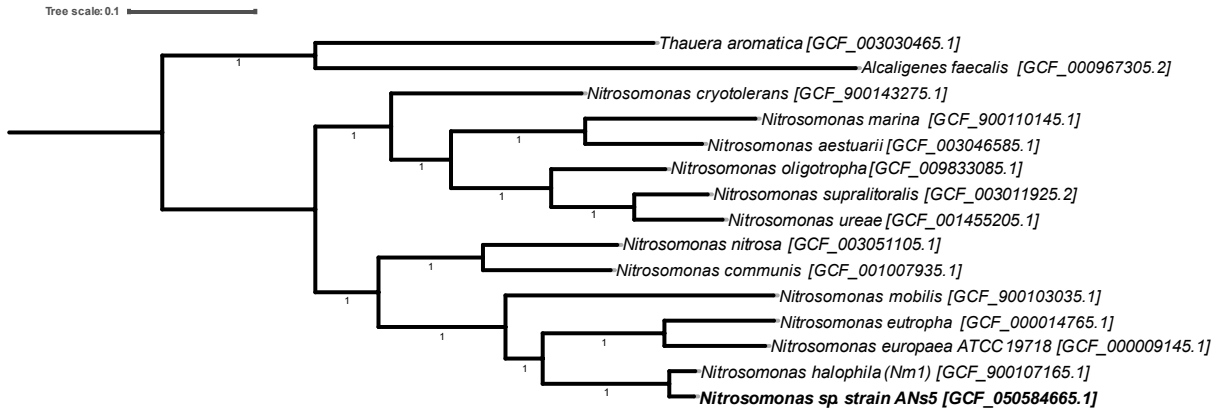

Figure 1 A maximum-likelihood phylogenetic tree of representative ammonia-oxidizing bacteria genomes visualized with iTOL (v 7.2) (13). The UFBoot support values are indicated below branches. The phylogenetic tree is rooted, with the outgroup represented by betaproteobacteria outside of the genus *Nitrosomonas*.

Data availability statement

The genome assembly is deposited at DDBJ/ENA/GenBank under BioProject PRJNA1125436. The raw sequencing data were deposited at NCBI SRA under accession number SRR29456553 and the assembly under accession number GCA_050584665.1.

Acknowledgments

JRV was supported by a United States National Science Foundation Graduate Research Fellowship (NSF-GRFP) and NSF award OCE-1924512 to AES. Sequencing was supported by NSF award ITE-2230641 to AES. J. Albers, P. Thieringer, and M. Li assisted with data analysis.

References

1. Sorokin D, Tourova T, Schmid M, Wagner M, Koops H-P, Kuenen G, Jetten M. 2001. Isolation and properties of obligately chemolithoautotrophic and extremely alkali-tolerant ammonia-oxidizing bacteria from Mongolian soda lakes. Arch Microbiol 176:170–177.
2. Koops H-P, Böttcher B, Böttcher B, Möller UC, Pommerening-Röser A, Stehr G. 1991. Classification of eight new species of ammonia-oxidizing bacteria: Nitrosomonas communis sp. nov., Nitrosomonas ureae sp. nov., Nitrosomonas aestuarii sp. nov., Nitrosomonas marina sp. nov., Nitrosomonas nitrosa sp. nov., Nitrosomonas eutropha sp. nov., Nitrosomonas oligotropha sp. nov. and Nitrosomonas halophila sp. nov. Microbiology 137:1689–1699.
3. Santoro AE, Casciotti KL, Francis CA. 2010. Activity, abundance and diversity of nitrifying archaea and bacteria in the central California Current: Nitrification in the central California Current. Environ Microbiol 12:1989–2006.
4. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. KBase: The United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 36:566–569.
5. Wingett SW, Andrews S. 2018. FastQ Screen: A tool for multi-genome mapping and quality control. F1000Research 7:1338.
6. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.
7. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863–864.
8. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol 13:e1005595.
9. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624.

10. Lee MD. 2019. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics 35:4162–4164.
11. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLOS ONE 5:e9490.
12. Letunic I, Bork P. 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res 52:W78–W82.
13. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. 2023. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. Nat Methods 20:1203–1212.
14. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114.