Complete genome sequences of two phylogenetically distinct *Nitrospina* isolated from the Atlantic and Pacific Oceans

Barbara Bayer[1#*], Matthew Kellom[2], Frederica Valois[3], John B. Waterbury[3], and Alyson E. Santoro[1*]

[1] Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, CA, USA

[2] DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

[3] Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

[#]Current address: Department of Microbiology and Ecosystem Science, University of Vienna, Vienna, Austria

*Correspondence: bbayer@ucsb.edu or asantoro@ucsb.edu

Keywords: Nitrospina, nitrite oxidation, nitrification, siderophore biosynthesis~~, DNA methylation~~

Abstract

The complete genome sequences of two chemoautotrophic nitrite-oxidizing bacteria of the genus *Nitrospina* ~~genus~~ are reported. *Nitrospina gracilis* strain Nb-211 was isolated from the Atlantic Ocean and *Nitrospina* sp. strain Nb-3 was isolated from the Pacific Ocean. We report two highly similar ~3.07 Mbp genome sequences that differ ~~in their DNA methylation patterns and~~by ~~in~~ the presence of ferric iron chelator (siderophore) biosynthesis ~~and transport~~ genes.

Announcement

Nitrospina are aerobic, chemoautotropic nitrite-oxidizing bacteria that have so far only been found in marine habitats (1), where they play an important role in the nitrogen cycle (2). *Nitrospina gracilis* Nb-211 (ATCC# 25379), first described in 1971 (3), was isolated from surface waters (13 m depth) of the Atlantic Ocean approximately 200 miles from the mouth of the Amazon R~~r~~iver ~~(0.1°25' N; 49°15'W)~~. *Nitrospina* sp. strain Nb-3 was isolated from the Pacific Ocean off the coast of Peru and has not been validly described ~~yet~~, however, its 16S rRNA gene sequence was published in 1994 (4). Both strains belong to the *Nitrospinaceae* family within the Nitrospinae/Nitrospinota (JGI/GTDB) phylum. ~~Both strains have been maintained in continuous liquid culture since their original isolation.~~

For genomic sequencing, cultures ~~Cells for genomic sequencing~~ were ~~cultivated~~ grown in 2 L glass bottles in artificial seawater medium containing 2 mM nitrite and ~~bottles were~~ incubated in the dark without agitation as described previously (5). Cells were collected via centrifugation (1 h, 15,000 *g*, 10°C) and ~~Genomic~~ DNA was extracted from cell pellets using a CTAB-Phenol-Chloroform protocol (6) ~~and~~. D~~d~~raft genomes were generated at the DOE Joint Genome Institute (JGI) using the Pacific Biosciences (PacBio) sequencing technology (7). Genomic DNA was sheared to 10 kb using g-TUBE columns (Covaris) and subjected to library preparation using the SMRTbell Express Template Prep 2.0 Kit. The PacBio SMRTbell library was purified and size-selected using AMPure PB beads and sequenced on the PacBio Sequel platform~~A >10 kb PacBio SMRTbell™ library was constructed and sequenced on the PacBio Sequel platform~~, which generated 123,391 ~~filtered~~ subreads (5,002.7 ± 3,337.2 bp) totaling 617,291,847 bp for strain Nb-211 ~~(JGI Sequencing Project ID: 1284499, Analysis Project ID: 1284485)~~ and 89,053 ~~filtered~~ subreads (6,935.6 ± 5,348.8 bp) totaling 617,633,853 bp for strain Nb-3 ~~(JGI Sequencing Project ID: 1284498, Analysis Project ID: 1284488), respectively~~. Reads >5 kb were assembled with HGAP (smrtlink/8.0.0.80529, HGAP 4 (1.0)) using default settings (8). The input read coverage was 188.4X for strain Nb-211 and 189.5X for strain Nb-3. The final draft genome sequences consisted of one scaffold each, with a total size of 3,069,626 bp ~~and a G+C content of 57.43%~~ for strain Nb-211~~,~~ and ~~a total size of~~ 3,075,869 bp ~~and a G+C content of 56.21%~~ for strain Nb-3 (Table 1). We confirmed complete circularization with the Circlator pipeline (v.1.5.5) (9), which uses nucmer (v.3.1) (10) to check for alignment between assembled contigs at opposite ends of the assembly, identifying a~~.~~ ~~The strain Nb-211 assembly scaffold end aligned and overlapped with the scaffold start over a length of~~ 50,007 bp alignment with 100% identity for strain Nb-211 and ~~.~~ ~~The strain Nb-3 assembly scaffold end aligned and overlapped with the scaffold start for a length of~~ 50,012 bp with 99.99% identity for strain Nb-3~~.~~

~~DNA modification detection and motif analysis was performed using the PacBio SMRT analysis platform (cromwell.workflows.pb_basemods) (14). Briefly, raw reads were filtered using SFilter, to remove short reads and reads derived from sequencing adapters. Filtered reads were then aligned to the reference genomes for strain Nb-211 and strain Nb-3 using BLASR(5.3) (15). Modified sites were identified through kinetic analysis of the aligned DNA sequence data (16) and grouped into motifs using MotifFinder (14). These motifs represent the recognition sequences of methyltransferase genes active in the genome (17) and differed between strain Nb-211 and Nb-3 (Table 1).~~

Table 1. Genome features of *Nitrospina gracilis* Nb-211 and *Nitrospina* sp. Nb-3

	Nitrospina gracilis Nb-211	*Nitrospina* sp. Nb-3

Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Italic, Font color: Auto
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt, English (United States)
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Helvetica, 11 pt
Formatted: Font: Italic
Formatted: Font: Italic
Formatted Table

GenBank accession	JAKJKD010000001	JAKJKC010000001
JGI Taxon ID	2917506613	2929071401
Genome size (bp)	3,069,626	3,075,869
G+C content (%)	57.43	56.21
DNA scaffolds	1	1
Total genes	2,939	2,905
Protein coding genes	2,879	2,846
rRNA operons	1	1
tRNA genes	50	49
CRISPR count	2	0

Formatted: Font: Helvetica, 10 pt

Formatted: Font: Helvetica, 10 pt

Formatted: Font: Helvetica, 10 pt

Formatted: Font: Helvetica, 10 pt

Motif	Position	Modification type	Count	Modified (%)	Mean score	Mean IDP ratio	Mean coverage	Objective score
Nitrospina gracilis Nb-211								
TTCGAA	6	m6A	1972	99.9	570	4.4	600	1123024
CTGAAG / CTTCAG	5	m6A	1629	100	669	5.1	594	1090398
GCAM	1	modified base	55505	2.4	115	1.9	610	5202
Nitrospina sp. Nb-3								
CGGAGA	6	m6A	2149	99.9	670	4.8	635	1437763
VCGWCGSNY	3	modified base	3586	30.5	156	2.1	644	58291
GGGCCCV	3	modified base	932	20.8	157	2.2	626	7426

Both genomes were annotated using the IMG Annotation Pipeline (IMGAP) v.5.0.22/3. The genome of *Nitrospina gracilis* Nb-211 contains 2,939 coding DNA sequences (CDS) and 50 tRNAs (JGI Taxon ID: 2917506613). The genome of *Nitrospina sp.* Nb-3 contains 2,905 CDS and 49 tRNAs (JGI Taxon ID: 2929071401). Both genomes contain one rRNA operon. Strains Nb-211 and Nb-3 both belong to the *Nitrospinaceae* family within the order Nitrospinales of the class Nitrospinia within the Nitrospinae/Nitrospinota (JGI/GTDB) phylum. Both genome sequences share an average nucleotide identity (ANI) of 85.5%, well below the intraspecies threshold of 96.5% (11). Strain Nb-3 shares 99.98 % ANI with the previously

Formatted: Indent: First line: 0"

published draft genome of *Nitrospina gracilis* strain 3/211 (12) indicating that the latter likely derives from the culture originally designated as strain Nb-3 (4).

~~Comparative genome analyses using OrthoFinder (18) indicated that the Nb-211 and Nb-3 genomes share ~99.4 % of their orthologues. Differences between the genomes include the presence of different methyltransferases, consistent with the observed methylation motifs unique to both genomes.~~ Strain Nb-3 encodes a putative iron chelator (siderophore) biosynthesis gene cluster~~and transport genes~~, which ~~are~~ is absent in strain Nb-211, potentially reflecting adaptations to differences in iron availability in the respective ocean basins the strains were isolated from. Strain Nb-211 was isolated near the Amazon river, which is a source of iron to the Atlantic Ocean (13), while strain Nb-3 was isolated from the relatively iron-deplete North Pacific (14).

Data availability

The whole-genome shotgun sequencing project of *Nitrospina gracilis* Nb-211 has been deposited at DDBJ/ENA/GenBank under BioProject number PRJNA708439 and accession number JAKJKD000000000. The whole-genome shotgun sequencing project of *Nitrospina sp.* Nb-3 has been deposited at DDBJ/ENA/GenBank under BioProject number PRJNA783628 and accession number JAKJKC000000000. The NCBI Sequence Read Archive (SRA) accession numbers for the raw reads are SRR17430281 for strain Nb-211 and SRR17430190 for strain Nb-3.

Acknowledgements

This work was supported by a Simons Foundation Early Career Investigator Award (3435889) and U.S. National Science Foundation award OCE-1924512 to AES. BB was supported by the Austrian Science Fund (FWF) Project Number J4426-B. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-05CH11231. These data were generated for JGI Proposal #506203 to BB and AES. We thank Christa Pennacchio who managed the project at JGI.

References

1. Daims H, Lücker S, Wagner M. 2016. A New Perspective on Microbes Formerly Known as Nitrite-Oxidizing Bacteria. Trends in Microbiology 24:699–712.

2. Kuypers MMM, Marchant HK, Kartal B. 2018. The microbial nitrogen-cycling network. Nature Reviews Microbiology 16:263–276.
3. Watson SW, Waterbury BJ. 1971. Characteristics of Two Marine Nitrite Oxidizing Bacteria, Nitrospina gracilis nov. gen. nov. sp. and Nitrococcus mobilis nov. gen. nov. sp. Archives of Microbiology 77:203–230.
4. Teske A, Alm E, Regan JM, Toze S, Rittmann BE, Stahl DA. 1994. Evolutionary Relationships among Ammonia- and Nitrite-Oxidizing Bacteria 176:6623–6630.
5. Bayer B, Saito MA, McIlvin MR, Lücker S, Moran DM, Lankiewicz TS, Dupont CL, Santoro AE. 2021. Metabolic versatility of the nitrite-oxidizing bacterium Nitrospira marina and its proteomic response to oxygen-limited conditions. ISME Journal 15:1025–1039.
6. CTAB-Phenol-Chloroform DNA extraction protocol: https://jgi.doe.gov/wp-content/uploads/2014/02/JGI-Bacterial-DNA-isolation-CTAB-Protocol-2012.pdf.
7. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, deWinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. 2009. Real-Time DNA Sequencing from Single Polymerase Molecules. Science 323:133 LP – 138.
8. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nature Methods 10:563–569.
9. Hunt M, Silva N De, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: Automated circularization of genome assemblies using long sequencing reads. Genome Biology 16:~~1–10~~294.
10. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. Genome biology 5:R12.
11. Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, Kyrpides NC, Pati A. 2015. Microbial species delineation using whole genome sequences. Nucleic Acids Research 43:6761–6771.
12. Lücker S, Nowka B, Rattei T, Spieck E, Daims H. 2013. The genome of Nitrospina gracilis illuminates the metabolism and evolution of the major marine nitrite oxidizer. Frontiers in Microbiology 4:27.

13. Subramaniam A, Yager PL, Carpenter EJ, Mahaffey C, Björkman K, Cooley S, Kustka AB, Montoya JP, Sañudo-Wilhelmy SA, Shipe R, Capone DG. 2008. Amazon River enhances diazotrophy and carbon sequestration in the tropical North Atlantic Ocean. Proceedings of the National Academy of Sciences 105:10460–10465.
14. Moore JK, Braucher O. 2007. Observations of dissolved iron concentrations in the World Ocean: implications and constraints for ocean biogeochemical models. Biogeosciences 4:1241–1277.