Supplementary Material

Genome-enabled insights into the ecophysiology of the comammox bacterium *Candidatus* Nitrospira nitrosa

Pamela Y. Camejo[1], Jorge Santo Domingo[2], Katherine D. McMahon[1,3], Daniel R. Noguera[1,*]

[1]Department of Civil and Environmental Engineering, University of Wisconsin - Madison, Madison, WI, USA

[2] Environmental Protection Agency, Cincinnati, OH, USA

[3]Department of Bacteriology, University of Wisconsin - Madison, Madison, WI, USA

Supplementary Tables

Table S1. Metrics of the two *Nitrospira* genomes assembled in this study after each genome refinement step.

Bin Id	Assembly Step	GC (%)	Genome size (Mbp)	Gene count	# Scaffolds	N50	Completeness (%)	Contamination (%)	Strain heterogeneity (%)
UW-LDO-01	MMgenome	55.1	4.13	4479	523	17,227	95.9	3.79	0
	Contigs length >1,000	55.0	4.08	4330	422	17,431	95.9	3.79	0
	SSpace Scaffolding	55.0	4.09	4268	287	27,733	95.9	3.74	0
	SSpace Gap Filling	55.0	4.08	4247	287	27,614	95.9	3.74	0
	Removal of scaffolds with non-*Nitrospira* ORFs	54.9	3.91	4031	230	29,398	95.9	3.64	0
UW-LDO-02	MMgenome	59.3	4.17	5360	1,878	4,227	97.5	15.4	0
	Contigs length >1,000	59.2	3.69	4262	1,003	4,891	94.3	6.10	0
	SSpace Scaffolding	59.2	3.72	4082	584	10,904	94.3	4.66	0
	SSpace Gap Filling	59.2	3.70	3993	584	10,895	94.4	4.66	0
	Removal of scaffolds with non-*Nitrospira* ORFs	59.2	3.54	3809	524	11,000	94.4	4.66	0

Completeness, contamination and strain heterogeneity were assessed with CHECKM 0.7.1, according to the presence of 43 single-copy reference genes.

Table S2. Genomes included in the metagenomic mapping analysis and the number of reads mapping to each of them. The number of mapping reads was also normalized by the number of reads in each metagenome, paired-end reads average length and genome size.

			Metagenomic Sample							
			Day	100	**Day**	317	**Day**	522	**Day**	674
			Reads Length	125	**Reads Length**	125	**Reads Length**	150	**Reads Length**	125
			Nº Reads	12,630,882	**Nº Reads**	16,037,438	**Nº Reads**	109,197,498	**Nº Reads**	16,056,492
Functional Group	**Genome**	**Length (bp)**	**Mapping Reads**	**Normalized Counts**	**Mapping Reads**	**Normalized Counts**	**Mapping Reads**	**Normalized Counts**	**Mapping Reads**	**Normalized Counts**
AOA	*Ca.* Nitrosoarchaeum limnia BG20 (AHJG01000000)	1,855,559	4	0	8	0	0	0	0	0
	Ca. Nitrosoarchaeum limnia SFB1 (CM001158)	1,772,718	2	0	2	0	0	0	2	0
	Nitrososphaera viennensis EN76 (CP007536)	2,527,938	38	0	48	0	0	0	36	0
	Cenarchaeum symbiosum A (DP000238)	2,045,086	38	0	40	0	0	0	19	0
	Nitrosopumilus maritimus SCM1 (CP000866)	2,673,040	4	0	2	0	0	0	1	0
	Nitrosopumilus sp. AR (AJVJ01000001)	4,084,671	-	-	6	0	0	0	2	0
AOB	*Nitrosococcus oceani* (NC_007484)	3,481,691	992	3	2,032	5	21,760	9	964	2
	Nitrosococcus mobilis (FMWO00000000)	3,082,864	8,334	27	1,834	5	4,690	2	259	1
	Nitrosomonas communis Nm2 (NZ_CP011451)	4,067,838	686	2	252	0	982	0	66	0
	Nitrosomonas europaea (AL954747)	2,812,094	2,702	10	3,138	9	10,724	5	586	2
	Nitrosomonas eutropha (NC_008344)	2,661,057	842	3	316	1	1,760	1	96	0
	Nitrosomonas sp. AL212 (NC_015222)	3,180,526	910	3	340	1	6,868	3	287	1
	Nitrosomonas sp. Is79A3 (CP002876)	3,783,444	2,678	7	1,346	3	10,166	4	330	1
	Nitrosomonas ureae Nm10 (CP013341)	3,307,820	810	2	214	1	2,512	1	99	0
	Nitrosomonas aestuarii (FOSP01000000)	3,994,034	594	1	434	1	1,144	0	90	0
	Nitrosomonas halophile (FNOY01000000)	1,645,259	880	5	610	3	1,366	1	167	1
	Nitrosomonas nitrosa (FOUF01000000)	3,154,865	400	1	268	1	808	0	51	0
	Nitrosospira briensis (NZ_CP012371)	3,441,569	630	2	432	1	984	0	146	0
	Nitrosospira multiformis (CP000103)	3,184,243	614	2	408	1	844	0	111	0
	Nitrosospira sp. NpAV (JXQM01000000)	3,210,113	1,096	3	464	1	1,108	0	150	0
Comammox	*Nitrospira* sp. UW-LDO-01	3,909,139	545,070	1,380	244	0	990	0	105	0
	Nitrospira sp. Ga0074138 (LNDU01000000)	4,111,525	2,980	7	302	1	26	0	71	0
	Ca. Nitrospira nitrosa (CZQA01000000)	4,422,398	41,490	93	188	0	88	0	48	0
	Ca. Nitrospira nitrificans (CZPZ01000000)	2,385,067	4,766	20	410	1	86	0	108	0
	Ca. Nitrospira inopinata (LN885086)	3,155,724	1,628	5	360	1	278	0	125	0
NOB	*Nitrospira moscoviensis* NSP_M-1 (NZ_CP011801)	4,589,485	1,852	4	644	1	82	0	253	0
	Nitrospira_defluvii (NC_014355)	4,317,083	395,170	906	113,576	205	8,268	3	38,506	69
	Nitrospira sp. UW-LDO-02	3,541,761	771,620	2,157	101,916	224	4,562	2	18,577	41
	Nitrobacter winogradskyi (NC_007406)	3,402,093	902	3	1,066	2	9,950	4	759	2
	Nitrobacter hamburgensis (NC_007964)	4,406,967	1,784	4	1,744	3	16,432	5	1,018	2
	Nitrospira sp. OLB23 (JZQY00000000)	3,754,263	6,276	17	2,086	4	2,218	1	516	1
Anammox	Ca. Brocadia fulgida (LAQJ01000000)	3,552,939	36	0	40	0	78	0	21	0
	Ca. Brocadia caroliniensis (AYTS01000000)	3,728,197	20	0	42	0	6	0	16	0
	Ca. Kuenenia stuttgartiensis (AMCF01000000)	3,810,257	154	0	148	0	46	0	339	1
	Ca. Brocadia sinica (BAFN01000000)	4,077,002	28	0	32	0	46	0	15	0
	Ca. Jettenia caeni (BAFH01000000)	4,087,350	20	0	32	0	22	0	12	0
		Total	**1,796,050**	**4,667**	**234,730**	**476**	**108,696**	**43**	**63,548**	**126**

Table S3. Inventory of genes involved in energy-driving processes from complete and draft genomes of *Nitrospira*. Grey and white rectangles represent presence and absence of each gene, respectively.

			Comammox					**NOB**			
Category	**Gene**	**Function**	***Nitrospira* sp. UW-LDO-01**	***Ca.* Nitrospira nitrosa**	***Ca.* Nitrospira nitrificans**	***Ca.* Nitrospira inopinata**	***Nitrospira* sp. Ga0074138**	***Nitrospira* sp. UW-LDO-02**	***Nitrospira moscoviensis***	***Nitrospira defluvii***	***Nitrospira* sp. OL23**
Ammonia Oxidation	***amoA***	**Ammonia monooxygenase subunit alpha**	■	■	■	■	■	□	□	□	□
	amoB	Ammonia monooxygenase subunit beta	■	■	■	■	■	□	□	□	□
	amoC	Ammonia monooxygenase subunit gamma	■	■	■	■	■	□	□	□	□
Hydroxylamine Oxidation	*haoA*	Hydroxylamine oxidoreductase subunit alpha	■	■	■	■	■	□	□	□	□
	haoB	Hydroxylamine oxidoreductase subunit beta	■	■	■	■	■	□	□	□	□
	cycA	Cytochrome c protein c554	■	■	■	■	■	□	□	□	□
	cycB	Cytochrome c protein, cM552	■	■	■	■	■	□	□	□	□
Nitrite Oxidation	*nxrA*	Nitrite oxidoreductase, alpha subunit	■	■	■	■	■	■	■	■	■
	nxrB	Nitrite oxidoreductase, beta subunit	■	■	■	■	□	□	■	■	■
Nitrite Reduction	*nirK*	Copper- containing dissimilatory nitrite reductases	■	■	■	■	■	■	■	■	■
	nrfA	cytochrome c nitrite reductase	□	□	□	■	□	□	□	□	□
Urea Hydrolysis	*urtA*	Urea ABC transporter	■	■	■	■	■	□	■	□	□
	urtB	Urea ABC transporter	■	■	■	■	■	□	□	□	□
	urtC	Urea ABC transporter	■	■	■	■	■	□	□	□	□
	urtD	Urea ABC transporter	■	■	■	■	■	□	□	□	□
	urtE	Urea ABC transporter	■	■	■	■	■	□	□	□	□
	ureA	Urease subunit gamma	■	■	■	■	■	□	■	□	□
	ureB	Urease subunit beta	■	■	■	■	■	□	■	□	□
	ureC	Urease subunit alpha	■	■	■	■	■	□	■	□	□
	ureD	Urease accessory protein	■	■	■	■	■	□	■	□	□
	ureF	Urease accessory protein	■	■	■	■	■	□	■	□	□
	ureG	Urease accessory protein	■	■	■	■	■	□	■	□	□
Cyanase degradation	*cynS*	Cyanase hydratase	□	□	□	□	□	■	■	■	■
Formate Oxidation	*fdsA*	Formate dehydrogenase subunit alpha	□	□	□	□	□	■	■	■	■
	fdsB	Formate dehydrogenase subunit beta	□	□	□	□	□	■	■	■	■
	fdsC	Formate dehydrogenase subunit gamma	□	□	□	□	□	■	■	■	■
	focA	Formate transporter	□	□	□	□	□	■	■	■	■
Hydrogen Oxidation	*hoxA*	Hydrogenase transcriptional regulatory protein	□	□	□	□	□	□	■	□	□
	hupS	Putative [NiFe] hydrogenase small subunit	□	□	□	□	□	□	■	□	□
	hupL	Putative [NiFe] hydrogenase large subunit	□	□	□	□	□	□	■	□	□
	hypA	Hydrogenase nickel incorporation protein	■	■	■	■	■	□	■	□	□
	hypB	Hydrogenase nickel incorporation protein	■	■	■	■	■	□	■	□	□
	hypC	Hydrogenase expression/formation protein	■	■	■	■	■	□	■	□	□
	hypD	Hydrogenase expression/formation protein	■	■	■	■	■	□	■	□	□
	hypE	Hydrogenase maturation protein	■	■	■	■	■	□	■	□	□
	hypF	Hydrogenase metallocenter assembly protein	□	□	□	□	□	□	■	□	□
	hyfB	Hydrogenase, membrane subunit HyfB	■	■	□	□	■	■	■	■	■
	hyfC	Hydrogenase, membrane subunit HyfC	■	■	□	□	■	■	■	■	■
	hyfE	Hydrogenase, membrane subunit HyfE	■	■	□	□	■	■	■	■	■
	hyfF	Hydrogenase, membrane subunit HyfF	■	■	□	□	■	■	■	■	■
	hyfG	Hydrogenase, large subunit HyfG	■	■	□	□	■	■	■	■	■
	hyfI	Hydrogenase, small subunit HyfI	■	■	□	□	■	□	■	■	■
Sulfur reduction	*hydA*	Sulfhydrogenase, subunit alpha	■	■	■	■	■	□	□	□	□
	hydB	Sulfhydrogenase, subunit beta	■	■	■	■	■	□	□	□	□
	hydD	Sulfhydrogenase, subunit delta	■	■	■	■	■	□	□	□	□
	hydG	Sulfhydrogenase, subunit gamma	■	■	■	■	■	□	□	□	□
	hybD	Hydrogenase maturation preotease	■	■	■	■	■	□	□	□	□

Supplementary Figures

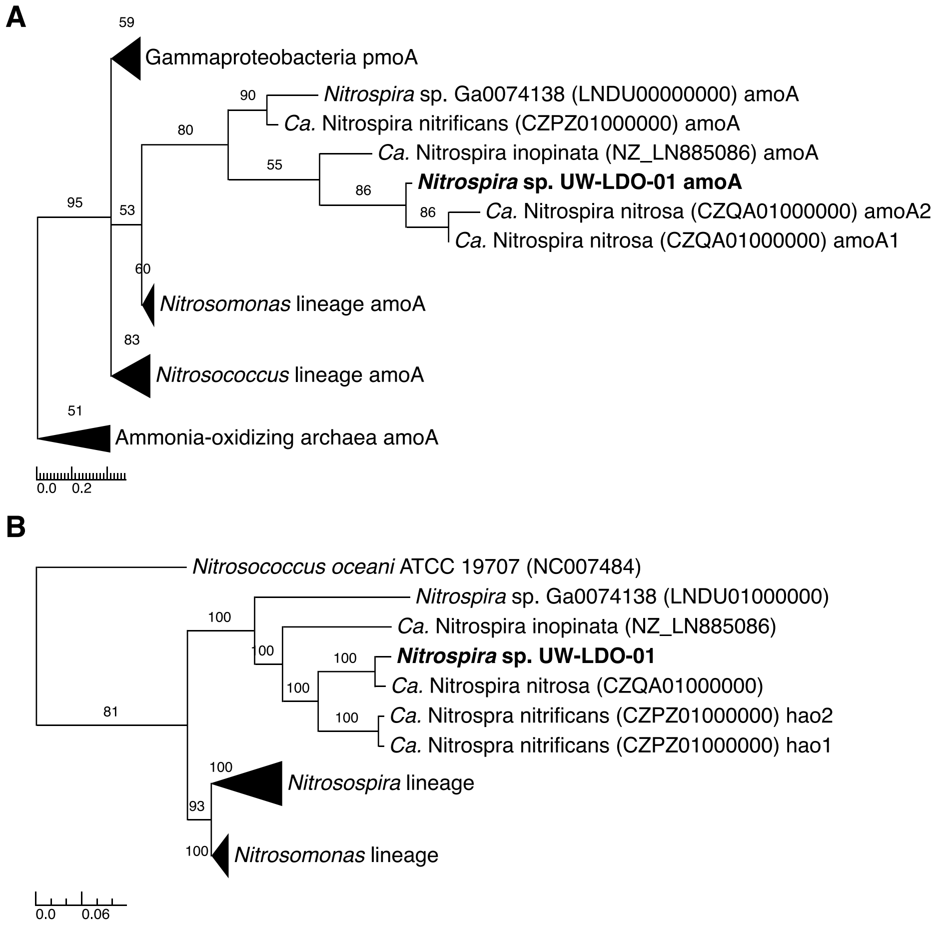

Figure S1. Neighbor-joining phylogenetic tree based on nucleic acid sequences of homologs to (A) *amoA* found in the genome of ammonia oxidizing microorganisms, methane oxidizing bacteria and comammox and (B) *hao* found in the genome of ammonia oxidizing bacteria and comammox. Bootstrap values are shown in the tree branches based on 1000 bootstrap replicates. The scale bar represents the number of nucleotide substitutions per site.

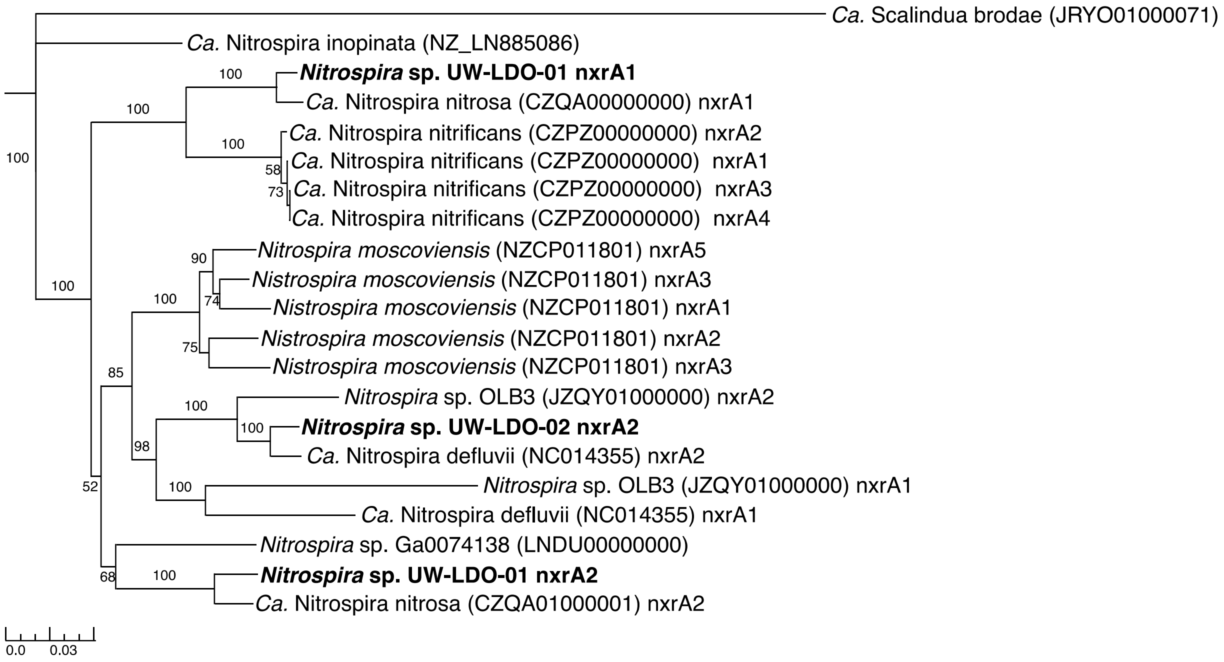

Figure S2. Neighbor-joining phylogenetic tree based on nucleic acid sequences of homologs to *nxr* found in the genome of *Nitrospira*-like bacteria. Bootstrap values are shown in the tree branches based on 1000 bootstrap replicates. The scale bar represents the number of nucleotide substitutions per site.