



Metagenomes from 25 Low-Abundance Microbes in a Partial Nitrification Anammox Microbiome

 Natalie K. Beach,^{a,b}  Kevin S. Myers,^{c,d}  Timothy J. Donohue,^{c,d,e}  Daniel R. Noguera^{a,c,d}

^aDepartment of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, Wisconsin, USA

^bCarollo Engineers, Inc., Broomfield, Colorado, USA

^cGreat Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, Wisconsin, USA

^dWisconsin Energy Institute, University of Wisconsin-Madison, Madison, Wisconsin, USA

^eDepartment of Bacteriology, University of Wisconsin-Madison, Madison, Wisconsin, USA

Natalie K. Beach and Kevin S. Myers contributed equally to this work. Author order was determined alphabetically.

ABSTRACT Microbial communities using anammox bacteria to remove nitrogen are increasingly important in wastewater treatment. We report on 25 metagenome-assembled genomes of low-abundance microbes from a partial nitrification anammox bioreactor system that have not been described previously. These data add to the body of information about this important wastewater treatment system.

Processes utilizing anammox bacteria for nitrogen removal have become popular for energy-saving wastewater treatment since they do not require oxygen or organic carbon for growth to convert ammonium (NH_4^+) and nitrite (NO_2^-) directly to nitrogen gas (N_2) (1–7). Microorganisms in partial nitrification anammox (PNA) bioreactors adapt to repetitive and rapidly changing microaerobic and anoxic environments (5). The most abundant and active functional groups involved in nitrogen cycling in PNA systems are anammox bacteria, ammonia-oxidizing bacteria, and nitrite-oxidizing bacteria, and some of the strains appear to be ubiquitous (8, 9); we are beginning to realize the extent of the metabolic versatility and interactions within these microbiomes (8–10). We reported on metagenome-assembled genomes (MAGs) of the most abundant and active microorganisms in a laboratory-scale PNA bioreactor treating reject water from a full-scale struvite recovery process (11). There, the bioreactor was inoculated with biomass from the York River Treatment Plant (Hampton Roads Sewerage District, Seaford, VA), which uses a PNA deammonification process to treat reject water from the solids dewatering facility (12).

This announcement includes 25 low-abundance MAGs with greater than 68% completion, grouped into 16 clusters, that were recovered from the same laboratory-scale PNA bioreactor and for which a specific role in the community has not been described (Table 1). These additional MAGs add to the expanding body of knowledge about microorganisms present in deammonification bioreactors at low abundance (e.g., each less than 1% based on DNA sequencing coverage and collectively less than 5%, determined as described previously [11, 13]).

Genomic DNA was extracted at multiple time points during operation of the bioreactor (days 77, 231, 350, and 454) using a modified phenol-chloroform method (11). The quality of the isolated DNA was determined using a Qubit 4 fluorometer (Thermo Fisher Scientific, MA, USA), a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific), and gel electrophoresis. Sequencing libraries were prepared using the TruSeq DNA PCR-free kit (Illumina, CA, USA) following the standard protocol as described (14), samples were sequenced on the HiSeq 2500 platform (Illumina) at the University of Wisconsin-Madison Biotechnology Center using 2×250 -bp reads, and low-quality reads (quality scores of less than 20 and sequence lengths of less than 100 bp) were removed using Sickel (15). A total of 18,106,239 reads

Editor Irene L. G. Newton, Indiana University, Bloomington

Copyright © 2022 Beach 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 Daniel R. Noguera, noguera@engr.wisc.edu.

The authors declare no conflict of interest.

Received 2 March 2022

Accepted 29 April 2022

TABLE 1 MAG statistics and genome accession numbers

MAG identification	GenBank accession no.	GTDB taxonomy	Completeness (%)	Contamination (%)	Genome size (bp)	No. of scaffolds	N ₅₀ (bp)	GC content (%)	Sequencing depth (×)	No. of rRNAs	No. of 5S rRNAs	No. of 16S rRNAs	No. of 23S rRNAs
BCT_39	JAIVC0000000000	Bacteroidia bacterium	99.52	3.02	4,874,625	55	472,994	42.53	87	45	2	2	2
BAC_79R	JAIVK0000000000	Bacterium	98.88	4.49	5,858,271	32	291,821	61.52	15	52	0	0	1
GAM_9R	JAIVIO0000000000	Gammaproteobacteria bacterium	97.62	2.41	3,557,069	39	193,302	66.55	32	49	1	0	0
BAC_63R	JAIVF0000000000	Bacterium	97.22	0	3,260,764	4	1,880,758	62.82	17	48	1	0	1
TMB_89R	JAIVIN0000000000	Thermoanaerobaculla bacterium	97.01	0.85	6,170,929	69	148,422	65.82	16	48	1	1	1
BAC_38	JAIVB0000000000	Bacterium	95.54	1.1	7,631,068	45	347,644	58.95	126	48	1	1	1
ELM_99	JAIVP0000000000	Elusimicrobia bacterium	95.51	0	3,644,178	43	213,845	48.71	30	76	1	1	1
PCS_19R	JAIVHX0000000000	Physcisphaerae bacterium	95.45	1.14	3,568,264	28	220,153	55.42	21	48	0	1	1
OMN_62	JAIVIE0000000000	Omnitrophota bacterium	95.16	0	2,713,171	25	1,594,128	61.88	33	48	1	1	1
PCS_14R	JAIVHU0000000000	Physcisphaerae bacterium	94.89	1.14	4,668,787	70	104,852	65.74	52	49	0	1	1
VRM_23R	JAIVHY0000000000	Verrucomicrobiae bacterium	94.59	3.07	4,819,195	36	167,251	59.69	18	50	1	1	1
PCS_76R	JAIVII0000000000	Physcisphaerae bacterium	94.32	0.57	5,487,514	72	107,025	67.82	16	54	1	1	1
PCS_15R	JAIVH0000000000	Physcisphaerae bacterium	93.18	1.7	5,249,578	50	159,786	67.13	59	61	0	1	1
ANL_81R	JAIVHM0000000000	Anaerolineae bacterium	92.73	2.18	7,897,783	60	252,904	57.36	13	48	1	1	1
PSM_69	JAIVG0000000000	Pseudomonadales bacterium	91.16	0.86	3,834,276	66	58,562	67.83	10	0	0	0	0
GAM_33	JAIVHZ0000000000	Gammaproteobacteria bacterium	89.29	3.48	4,271,281	128	64,536	64.03	45	52	1	0	0
MYX_71R	JAIVH0000000000	Myxococcota bacterium	89.03	0.65	3,968,878	78	77,580	63.12	11	44	1	1	0
RDC_77	JAIVJ0000000000	Rhodocyclaceae bacterium	88.03	0.21	2,815,254	67	57,548	64.68	18	39	0	0	0
PLM_90R	JAIVHO0000000000	Planctomycetota bacterium	85.88	1.14	4,250,401	91	61,976	64.89	9	65	0	1	1
PYM_42R	JAVID0000000000	Pyrimonadales bacterium	85.47	1.71	6,468,668	221	38,096	57.8	90	72	1	1	1
ANL_34	JAIVIA0000000000	Anaerolineae bacterium	77.45	0.91	3,991,433	141	33,931	56.33	49	38	0	0	0
BAC_13	JAIVHT0000000000	Bacterium	76.31	0	4,995,932	231	22,649	53.05	44	36	1	1	1
PLM_16R	JAIVHW0000000000	Planctomycetota bacterium	75.57	2.84	3,864,190	174	25,696	72.1	42	51	1	1	1
BRK_7R	JAIVL0000000000	Burkholderiaceae bacterium	69.09	0.29	2,770,384	154	18,435	69.49	18	39	0	0	0

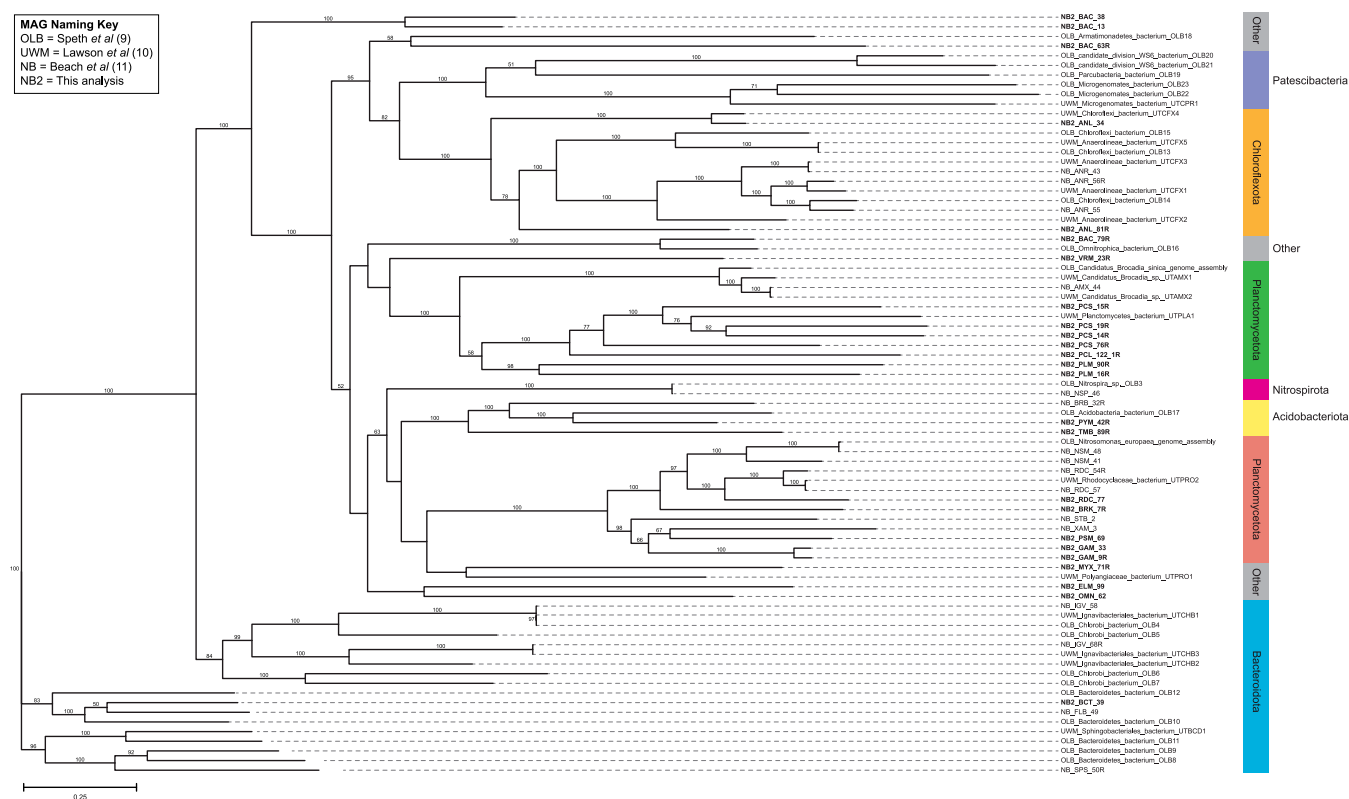


FIG 1 Phylogenetic tree of the 25 MAGs from this study (labeled NB2 and bolded), compared with the 16 most abundant MAGs from Beach *et al.* from the same microbiome (labeled NB) (11), as well as MAGs from Speth *et al.* (labeled OLB) (8) and Lawson *et al.* (labeled UWM) (9). The tree was constructed with RAXML-NG (unrooted) (25) with the 120 bacterial single-copy marker genes from GTDB-Tk (24) and visualized with TreeViewer (<https://treeviewer.org>). GTDB phylum taxonomy is listed on the right. Bootstrap values of more than 50 are shown. The scale bar indicates the number of nucleotide substitutions per sequence site. Strain codes for NB and NB2 strains are as follows: AMX, *Brocadia*; ANL, *Anaerolineae*; ANR, *Anaerolineae*; BAC, bacterium; BCT, *Bacteroidia*; BRB, *Bryobacteraceae*; BRK, *Burkholderiaceae*; ELM, *Elusimicrobia*; FLB, *Flavobacteriales*; GAM, *Gammaproteobacteria*; IGX, *Ignavibacteria*; MYX, *Myxococcota*; NSM, *Nitrosomonas*; NSP, *Nitrospira*; OMN, *Omnitrophota*; PCL, *Phycisphaerales*; PCS, *Phycisphaerae*; PLM, *Planctomycetota*; PSM, *Pseudomonadales*; PYM, *Pyrinomonadales*; RDC, *Rhodocyclaceae*; SPS, *Saprospiraceae*; STB, *Steroidobacteraceae*; TMB, *Thermoanaerobaculia*; VRM, *Verrucomicrobiae*; XAM, *Xanthomonadales*.

from the metagenomic samples were processed with default parameters as described previously (11), using SPAdes v3.3.0 (16) for assembly, Anvi'o v5.5.0 (17) for binning, BMap v38.22 (18) for mapping, Prodigal v2.6.3 (19), HMMER v3.2.1 (20), and NCBI Clusters of Orthologous Groups (COGs) (21) for annotation, and ProDeGe v2.2 (22) and tetranucleotide frequency comparisons for further cleaning of the MAGs. Genome statistics were determined with CheckM v1.0.3 (23), and taxonomy was assigned with GTDB-Tk v1.5.1 (database release 202) (24). The unrooted phylogenetic tree (Fig. 1) comparing these MAGs with MAGs from previous studies (8, 9) was generated using RAXML-NG v0.9.0 (25) and visualized with TreeViewer (<https://treeviewer.org>). These MAGs add to our understanding of PNA bioreactors and may aid in optimization of these systems once the function of these low-abundant microbes is elucidated.

Data availability. Raw metagenomic sequence data and MAGs are available in NCBI GenBank under BioProject accession number [PRJNA559529](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA559529). All custom scripts are available at GitHub (https://github.com/GLBRC/metagenome_analysis).

ACKNOWLEDGMENTS

This work was partially supported by funding from the National Science Foundation (grant CBET-1803055), the Madison Metropolitan Sewerage District (Madison, WI), and the Great Lakes Bioenergy Research Center (U.S. Department of Energy, Office of Science) (award DE-SC0018409).

REFERENCES

- van de Graaf AA, Mulder A, de Bruijn P, Jetten MS, Robertson LA, Kuenen JG. 1995. Anaerobic oxidation of ammonium is a biologically mediated process. *Appl Environ Microbiol* 61:1246–1251. <https://doi.org/10.1128/aem.61.4.1246-1251.1995>.

2. Kuenen JG. 2008. Anammox bacteria: from discovery to application. *Nat Rev Microbiol* 6:320–326. <https://doi.org/10.1038/nrmicro1857>.
3. Morales N, Val Del Rio A, Vazquez-Padin JR, Mendez R, Mosquera-Corral A, Campos JL. 2015. Integration of the Anammox process to the rejection water and main stream lines of WWTPs. *Chemosphere* 140:99–105. <https://doi.org/10.1016/j.chemosphere.2015.03.058>.
4. Wang S, Peng Y, Ma B, Wang S, Zhu G. 2015. Anaerobic ammonium oxidation in traditional municipal wastewater treatment plants with low-strength ammonium loading: widespread but overlooked. *Water Res* 84: 66–75. <https://doi.org/10.1016/j.watres.2015.07.005>.
5. Ma B, Wang S, Cao S, Miao Y, Jia F, Du R, Peng Y. 2016. Biological nitrogen removal from sewage via anammox: recent advances. *Bioresour Technol* 200:981–990. <https://doi.org/10.1016/j.biortech.2015.10.074>.
6. McCarty PL. 2018. What is the best biological process for nitrogen removal: when and why? *Environ Sci Technol* 52:3835–3841. <https://doi.org/10.1021/acs.est.7b05832>.
7. Gao H, Scherson YD, Wells GF. 2014. Towards energy neutral wastewater treatment: methodology and state of the art. *Environ Sci Process Impacts* 16:1223–1246. <https://doi.org/10.1039/c4em00069b>.
8. Speth DR, In 't Zandt MH, Guerrero-Cruz S, Dutilh BE, Jetten MS. 2016. Genome-based microbial ecology of anammox granules in a full-scale wastewater treatment system. *Nat Commun* 7:11172. <https://doi.org/10.1038/ncomms11172>.
9. Lawson CE, Wu S, Bhattacharjee AS, Hamilton JJ, McMahon KD, Goel R, Noguera DR. 2017. Metabolic network analysis reveals microbial community interactions in anammox granules. *Nat Commun* 8:15416. <https://doi.org/10.1038/ncomms15416>.
10. Kartal B, van Niftrik L, Keltjens JT, Op den Camp HJ, Jetten MS. 2012. Anammox–growth physiology, cell biology, and metabolism. *Adv Microb Physiol* 60: 211–262. <https://doi.org/10.1016/B978-0-12-398264-3.00003-6>.
11. Beach NK, Myers KS, Owen BR, Seib M, Donohue TJ, Noguera DR. 2021. Exploring the meta-regulon of the CRP/FNR family of global transcriptional regulators in a partial-nitrification anammox microbiome. *mSystems* 6:e0090621. <https://doi.org/10.1128/mSystems.00906-21>.
12. Nifong A, Nelson A, Johnson C, Bott CB. 2013. Performance of a full-scale sidestream DEMON® deammonification installation. *Proc Water Environ Fed* 2013:3686–3709. <https://doi.org/10.2175/193864713813685700>.
13. Beach NK. 2019. Toward energy self-sufficient wastewater treatment: insights into biological nutrient removal using low dissolved oxygen. PhD dissertation. University of Wisconsin-Madison, Madison, WI.
14. Camejo PY, Domingo JS, McMahon KD, Noguera DR. 2017. Genome-enabled insights into the ecophysiology of the comammox bacterium “*Candidatus Nitrospira nitrosa*.” *mSystems* 2:504. <https://doi.org/10.1128/mSystems.00059-17>.
15. Joshi N, Fass J. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (version 1.33). <https://github.com/najoshi/sickle>.
16. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834. <https://doi.org/10.1101/gr.213959.116>.
17. Eren AM, Esen OC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319. <https://doi.org/10.7717/peerj.1319>.
18. Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner. <https://jgi.doe.gov/data-and-tools/software-tools/bbtools/bb-tools-user-guide/bbmap-guide/>.
19. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
20. Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput Biol* 7: e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
21. Galperin MY, Makarova KS, Wolf YI, Koonin EV. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* 43:D261–D269. <https://doi.org/10.1093/nar/gku1223>.
22. Tennessen K, Andersen E, Clingenpeel S, Rinke C, Lundberg DS, Han J, Dangi JL, Ivanova N, Woyke T, Kyrpides N, Pati A. 2016. ProDeGe: a computational protocol for fully automated decontamination of genomes. *ISME J* 10:269–272. <https://doi.org/10.1038/ismej.2015.100>.
23. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
24. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36: 1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>.
25. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.