Microbiology Resource Announcements

| Applied and Industrial Microbiology | Announcement

Metagenomes and metagenome-assembled genomes from microbial communities in a biological nutrient removal plant operated at Hamptons Road Sanitation District (HRSD) with high and low dissolved oxygen conditions

Blaise M. Enuh,[1,2] Kevin S. Myers,[1,2] Charles Bott,[3] Stephanie Klaus,[3] Kester McCullough,[3] Lilian McIntosh,[3] Natalie Beach,[4] Michelle Young,[4] Timothy J. Donohue,[1,2,5] Daniel R. Noguera[1,2,6]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Aeration is a major cost at biological nutrient removal (BNR) plants. We report on microbial communities in a pilot-scale BNR system before and after a dissolved oxygen transition from 2.5 to 0.2 mg/L implemented over 18 months. Four PacBio metagenomes and 316 metagenome-assembled genomes are announced.

KEYWORDS metagenomics, microbial communities, wastewater treatment, dissolved oxygen

Reducing dissolved oxygen (DO) levels in biological nutrient removal (BNR) processes can maintain efficient nitrification and phosphorus removal while lowering energy consumption (1–3). Decreased DO causes microbial community adaptation, though the dynamics remain poorly understood (1, 4). We report metagenomes and metagenome-assembled genomes (MAGs) from a BNR pilot plant operated at the pilot testing facility of the Hamptons Road Sanitation District (HRSD) Virginia Initiative Pilot (VIP) treatment plant in Norfolk, VA (5). Grab samples of mixed liquor were collected from the end of the aeration basin of the treatment plant before and after an 18-month transition from high- (2.5 mg/L) to low-DO (0.2 mg/L) conditions. Upon collection, the mixed liquor samples were centrifuged; the cell pellets were frozen, and then shipped overnight to the laboratory.

Genomic DNA was extracted directly from the pellets using the DNeasy PowerSoil Kit (Qiagen, Germantown, MD), then quantified using a Qubit fluorometer (Fisher Scientific, Waltham, MA) and stored at −20°C until sequencing. DNA purity was measured on a NanoDrop One (Fisher Scientific), and concentration was measured with the Qubit dsDNA High-Sensitivity Kit (Fisher Scientific).

HiFi library preparation and sequencing at the University of Wisconsin-Madison Biotechnology Center (Madison, WI) followed workflow PN 102-166-600 Version 04 (Pacific Biosciences, Menlo Park, CA) with standard parameters. Library integrity was evaluated on the FemtoPulse System (Agilent, Santa Clara, CA). The library was quantified using the Qubit dsDNA High Sensitivity Kit and sequenced on a Sequel II using Sequel Polymerase Binding Kit 2.2 following the standard protocol (Pacific Biosciences). Across the four samples, there was an average of 1,770,838 reads, with a range from 439,484 to 3,127,946 reads. The average N50 of the reads was 8,664 bp (range of 7,020 bp to 10,234 bp). The Circular Consensus Sequence (CCS) reads were assembled utilizing either metaFLYE (v2.9-b1768) (6) and polished with racon (v1.4.20) (7) for the high-DO samples or metaMDBG (v0.3) (8) with a racon module (v1.4.20) (7) for the low-DO samples. The reads were mapped onto the assemblies using minimap2 (v2.22-r1101) (9). Binning

Editor Frank J. Stewart, Montana State University, Bozeman, Montana, USA

Address correspondence to Daniel R. Noguera, noguera@engr.wisc.edu.

The authors declare no conflict of interest.

Received 22 December 2025
Accepted 7 April 2026
Published 14 May 2026

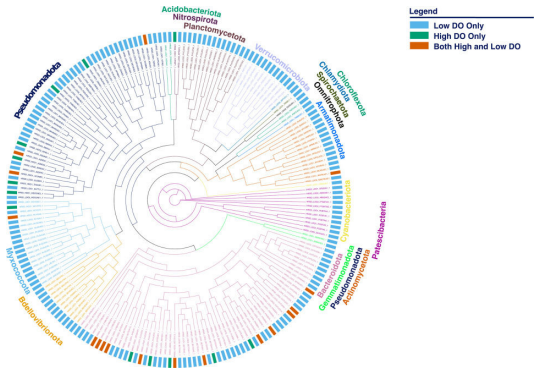

FIG 1 Phylogenetic map of the 207 unique MAG clusters from high and low-DO samples collected from the HRSD pilot plant. The outer ring is color-coded to indicate whether the cluster included MAGs from either low-DO, high-DO, or both conditions. MAG classification is indicated with names matching the color of tree branches.

was done with metaBAT2 (v2:2.15) (10). Contaminated contigs within each bin were identified with ProDeGe (v2.3) (11), and custom scripts were used for tetranucleotide frequency analysis (run.GC.sh, Calculating_TF_Correlations.R; https://github.com/GLBRC/metagenome_analysis). Contigs identified as contaminated with both methods were removed from the final MAG assemblies. CheckM (v1.2.2) (12) was used for quality evaluation, and taxonomy was determined using GTDB-Tk (v2.1.0) database release 09-RS214 (13). Functional annotations were assigned using Bakta (v1.9.1) (14). RAxML-NG adaptive (v1.2.1) (15) was used to infer the best phylogenetic tree by maximum likelihood estimation. The tree (Fig. 1) was visualized in TreeViewer (v2.2.0) (16) and annotated in Inkscape (v1.2.2) (https://inkscape.org). We obtained 316 MAGs from four metagenomes, with 278 MAGs from low-DO and 38 from high-DO. These MAGs were dereplicated with dRep (v0.6.1) (17) using a 95% identity cutoff, resulting in 207 unique MAGs with over 75% completeness and less than 10% contamination (Fig. 1 and in the Figshare dataset [18]). As treatment plants begin to embrace the concept of low-DO operation for energy savings (2, 5), these MAGs augment the knowledge of the microbial communities under low-DO operation.

The metagenomes, two from the high-DO and two from the low-DO condition, are available under BioProject number PRJNA1156690. The MAGs and their taxonomy can be accessed on Figshare (10.6084/m9.figshare.30389092 [18]). Custom scripts used in the analysis are available at https://github.com/GLBRC/metagenome_analysis.

ACKNOWLEDGMENTS

This work was supported by funding from the U.S. Department of Energy (DOE), Office of Energy Efficiency & Renewable Energy under award DE-EE0009509 and partially based upon work at the Great Lakes Bioenergy Research Center supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research under award number DE-SC0018409. DNA sequencing was done at the UW Biotechnology Center's DNA Sequencing Facility (Research Resource Identifier—RRID: SCR_017759).

AUTHOR AFFILIATIONS

[1]Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, Wisconsin, USA
[2]Wisconsin Energy Institute, University of Wisconsin-Madison, Madison, Wisconsin, USA
[3]Hamptons Road Sanitation District, Virginia Beach, Virginia, USA
[4]Carollo Engineers, Westminster, Colorado, USA
[5]Department of Bacteriology, University of Wisconsin-Madison, Madison, Wisconsin, USA

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

AUTHOR ORCIDs

Kevin S. Myers http://orcid.org/0000-0003-3302-3877
Natalie Beach http://orcid.org/0000-0002-4725-496X
Timothy J. Donohue http://orcid.org/0000-0001-8738-2467
Daniel R. Noguera http://orcid.org/0000-0003-0372-3063

AUTHOR CONTRIBUTIONS

Blaise M. Enuh, Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing – original draft | Kevin S. Myers, Data curation, Writing – review and editing | Charles Bott, Resources, Writing – review and editing | Stephanie Klaus, Resources, Writing – review and editing | Kester McCullough, Resources, Writing – review and editing | Natalie Beach, Conceptualization, Project administration, Resources, Writing – review and editing | Michelle Young, Conceptualization, Project administration, Resources, Writing – review and editing | Timothy J. Donohue, Conceptualization, Funding acquisition, Supervision, Writing – review and editing | Daniel R. Noguera, Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review and editing.

REFERENCES

1. Fitzgerald CM, Camejo P, Oshlag JZ, Noguera DR. 2015. Ammonia-oxidizing microbial communities in reactors with efficient nitrification at low-dissolved oxygen. Water Res 70:38–51. https://doi.org/10.1016/j.watres.2014.11.041
2. Keene NA, Reusser SR, Scarborough MJ, Grooms AL, Seib M, Santo Domingo J, Noguera DR. 2017. Pilot plant demonstration of stable and efficient high rate biological nutrient removal with low dissolved oxygen conditions. Water Res 121:72–85. https://doi.org/10.1016/j.watres.2017.05.029
3. Stewart RD, Bashar R, Amstadt C, Uribe-Santos GA, McMahon KD, Seib M, Noguera DR. 2022. Pilot-scale comparison of biological nutrient removal (BNR) using intermittent and continuous ammonia-based low dissolved oxygen aeration control systems. Water Sci Technol 85:578–590. https://doi.org/10.2166/wst.2021.630
4. Park HD, Noguera DR. 2004. Evaluating the effect of dissolved oxygen on ammonia-oxidizing bacterial communities in activated sludge. Water Res 38:3275–3286. https://doi.org/10.1016/j.watres.2004.04.047
5. Young M, Beach N, Reifsnyder S, Ackman P, Weiland T, Dreher B, Bott C, Ekster A, Enuh BM, Noguera D, Kestel S, Khatiwada S, McCollough K, McIntosh L, Nagarkar M, Rosso D, Rauch-Williams T. 2025. Transforming aeration energy in water resource recovery facilities (WRRFs) through suboxic nitrogen removal (final report). https://doi.org/10.2172/2587564.
6. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, Kuhn K, Yuan J, Polevikov E, Smith TPL, Pevzner PA. 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat Methods 17:1103–1110. https://doi.org/10.1038/s41592-020-00971-x
7. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate *de novo* genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116
8. Benoit G, Raguideau S, James R, Phillippy AM, Chikhi R, Quince C. 2024. High-quality metagenome assembly from long accurate reads with metaMDBG. Nat Biotechnol 42:1378–1383. https://doi.org/10.1038/s41587-023-01983-6
9. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191
10. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7:e7359. https://doi.org/10.7717/peerj.7359
11. Tennessen K, Andersen E, Clingenpeel S, Rinke C, Lundberg DS, Han J, Dangl JL, Ivanova NN, 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
12. 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
13. Chaumeil P-A, 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
14. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom 7:13. https://doi.org/10.1099/mgen.0.000685
15. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455. https://doi.org/10.1093/bioinformatics/btz305
16. Bianchini G, Sánchez-Baracaldo P. 2024. TreeViewer: flexible, modular software to visualise and manipulate phylogenetic trees. Ecol Evol 14:e10873. https://doi.org/10.1002/ece3.10873
17. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J 11:2864–2868. https://doi.org/10.1038/ismej.2017.126
18. Enuh B, Myers KS, Noguera DR. 2025. Metagenomes and metagenome-assembled genomes from microbial communities in a biological nutrient removal plant operated at hamptons road sanitation district (HRSD) with high and low dissolved oxygen conditions. Dataset. https://doi.org/10.6084/m9.figshare.30389092.