Supporting information

Differential kinetics of nitrogen oxides reduction leads to elevated N_2O production by a Denitrifying Polyphosphate Accumulating Organisms (DPAOs) enrichment

Han Gao[a], Xiaotian Zhao[b], Lei Zhou[b], Fabrizio Sabba[a] and George F. Wells[a,*]

[a]Department of Civil and Environmental Engineering, Northwestern University, Evanston, Illinois

[b]Master of Science in Biotechnology Program, McCormick School of Engineering, Northwestern University, Evanston, Illinois

*Corresponding author: George F. Wells, Email: george.wells@northwestern.edu, Ph: +1-847-491-3109, Fax: +1-847-491-4011

Supporting Information - Methods

Chemical analyses

Acetate and N_2O were measured using a GC-FID (Fisher-Thermo scientific, USA) with a TR-FFAP column. A 1 µL sample was injected into the GC with 50:1 split ratio with constant flow rate of 1 ml/min. Helium was used as carrier gas. The following temperature program was used: initial, 70 ºC for 1 min, 10 ºC/min up to 180ºC, and holding at 180 ºC for 6 min. Gas phase N_2O was measured by GC-ECD with a TG-BOND Q column. 100 µL of gas sample was injected with a split ratio of 50:1 and column flow of 5 ml/min using helium as carrier gas. An isothermal temperature program was used with oven temperature setpoint of 35ºC with backflush started 3 min after injection. The detector temperature was 250 ºC and N_2 was used as makeup gas.

PCR amplification for amplicon sequencing

For each sample, we prepared duplicate DNA extracts. A two-step PCR amplification and barcoding were performed using the Fluidigm Biomark multiplex PCR strategy as previously described (Griffin and Wells 2017). First, forward primer 515f (5'-GTG CCA GCM GCC GCG GTA A) and reverse primer 806r (5'-GGA CTA CHV GGG TWT CTA AT) were used to amplify the V4 region of the universal bacterial 16S rRNA gene (Caporaso et al. 2012). A 20 µL PCR reaction was performed per DNA extract, using 2× Epicentre Premix F PCR mastermix (Epicentre, USA), 3.5 U/µL Expand HiFidelity Taq (Roche Diagnostics, USA), 200 nM of each primer, and 1 µL gDNA. The following PCR amplification temperature program was used: 95 °C for 5 minutes, followed by 28 cycles of 95 °C for 30 s, 55 °C for 45 s, and 68 °C for 30 s with a final elongation step at 68°C for 7 minutes. Amplicons from replicate first PCR runs were pooled before the second PCR amplification. Fluidigm primers with sequencing adapters and a sample-specific barcode were used to perform the second PCR step. The PCR reaction was performed using 2× Accuprime Supermix (ThermoFisher, USA), 50 µM forward and reverse primers (Fluidigm, USA), and 1 µL of template from the combined first round PCR products using the following temperature program: 95 °C for 5 minutes, 95 °C for 30s, 60°C for 30s, and 68 °C for 30s in a total of 8 cycles. The resulting amplicons were then send to the University of Illionis Chicago DNA Service Facility for further purification with a Qiagen PCR purification kit and sequence on the Illumina MiSeq V2 platform.

16S rRNA gene amplicon sequencing data processing

After removing barcodes, filtering for low-quality sequences (sequences with more than 1 error per 100 bases were removed), and merging reverse and forward paired end sequences, USEARCH v8.1.1861 was used to remove singletons and chimeras and to select representative OTUs based on a 97% identity cutoff (Edgar 2013, Edgar et al. 2011). The Greengenes sequence database was employed for phylogenetic inference for representative sequences from each OTU in the Quantitative Insights Into Microbial Ecology (QIIME) platform (Caporaso et al. 2010, Edgar 2010). Representative OTUs were also aligned via the SILVA Incremental Aligner (SINA) database to identify '*Candidatus* Competibacter phasphatis' and *Defluviicocus*, candidate glycogen accumulating organisms (GAOs), since neither is included in the Greengenes database (Pruesse et al. 2012). Samples were rarefied to the lowest sequencing depth for further diversity analysis.

Quantitive PCR (qPCR)

All qPCR assays except *nirK* and clade II *nosZ* were performed on a CFX Connect thermocycler (Bio-rad, USA) in a 20 µL reaction volume containing 10 µL of 2× SsoAdvanced Universal SYBR Green Supermix (Bio-rad, USA), 0.5 µM of primers and 2 µL of diluted DNA extracts. To quantify the abundance of clade II *nosZ*, the iTaq Universal SYBR Green Supermix (Biorad, USA) was used. For the quantification of *nirK*, FailSafe Green premix was made by mixing 2× FailSafe premix F (Epicentre, USA) with 10,000× SYBR Green 1 nucleic acid stain (Invitrogen, USA). The qPCR reaction volume for *nirK* was 20 µL with 10 µL of FailSafe Green premix, 0.5 U/reaction Expand HiFidelity taq enzyme (Roche Diagnostics, USA), 0.5 µM of primers and 2 µL of diluted gDNA. Each qPCR reaction was performed in triplicate.

Supporting Information - Tables

Table S1. The relative abundance (%) and standard deviation of the top 15 bacterial taxa at the genus level in the denitrifying EBPR biomass, based on 16S rRNA gene sequencing. Taxa names are given at the lowest possible assignment based on the Greengenes database.

Phylum	Lowest level of taxonomic assignment *	Total biomass (%)		Size fractions (µm) (%)			
		Day 0-48	Day 490-600	<150	150-350	350-600	>600
Proteobacteria	f_*Xanthomonadaceae*	3.0±1.1	19.9±3.0	24.1±5.5	21.7±1.2	20.9±1.6	6.9±1.9
Proteobacteria	g_*Candidatus* Accumulibacter	2.2±0.9	13.5±4.2	9.6±2.4	19.1±9.2	19.3±4.0	31.2±2.1
Proteobacteria	g_*Candidatus* Competibacter	0.9±0.9	0.126±0.078	22.9±1.2	21.0±2.4	12.9±0.1	2.0±0.5
Chlorobi	p_*Chlorobi*	0.5±0.6	6.9±2.0	5.1±1.9	4.0±1.9	5.5±2.5	7.8±1.5
Proteobacteria	g_*Defluviicoccus*	0.3±0.1	9.2±6.0	2.3±0.9	1.7±0.2	1.3±0.5	0.4±0.0
GN02	p_GN02	0.2±0.1	1.9±3.0	2.6±3.0	3.7±4.6	6.9±8.4	13.0±11.0
Proteobacteria	f_*Comamonadaceae*	5.4±1.2	0.8±0.4	1.6±0.9	1.4±0.9	2.3±1.6	4.6±1.9
Chloroflexi	g_*Caldilinea*	0.1±0.1	2.4±0.3	4.9±2.6	3.5±2.5	1.9±0.9	0.5±0.2
Proteobacteria	f_*Comamonadaceae*	4.9±1.9	2.2±1.2	0.8±0.6	0.6±0.5	0.6±0.4	0.3±0.1
Proteobacteria	f_*Rhodocyclaceae*	2.4±0.9	1.7±1.3	1.0±0.3	0.7±0.2	1.1±0.7	1.9±1.4
Bacteroidetes	p_*Bacteroidetes*	0.0±0.0	1.9±0.4	0.7±0.4	0.5±0.3	1.2±0.0	5.3±0.9
Chloroflexi	f_*Anaerolineae*	0.2±0.1	1.3±0.3	1.4±0.7	1.6±1.0	2.2±1.5	2.2±0.9
Proteobacteria	o_*Rhizobiales*	0.6±0.4	0.7±0.4	1.9±1.0	2.1±1.5	2.0±1.3	1.4±0.8
Proteobacteria	f_*Phyllobacteriaceae*	0.2±0.2	1.1±0.4	1.8±0.8	1.8±1.2	1.6±0.9	0.7±0.2
Proteobacteria	g_*Luteimonas*	0.1±0.0	2.2±1.4	0.3±0.2	0.3±0.1	0.4±0.0	0.2±0.1

* Lowest level of taxonomic assignment (p: phylum, c: class, o: order, f: family and g: genus).

Table S2. qPCR primers and thermal cycling conditions used in this study.

Gene	Primers	Size	Step	Temperature	Time	Reference
Universal 16S	519f: CAG CMG CCG CGG TAA NWC 907r: CCG TCA ATT CMT TTR AGT T	393bp	Initial Denaturation	95ºC	5min	(Burgmann et al. 2011)
			30 cycles	95ºC 60ºC 68ºC	30s 30s 30s	
Accumulibacter 16S	518f: CCA GCA GCC GCG GTA AT 846r: GTT AGC TAC GGC ACT AAA AGG	351bp	Initial Denaturation	95ºC	3min	(He et al. 2007)
			35 cycles	95ºC 59ºC 72ºC	30s 60s 30s	
narG	narG-f TCG CCS ATY CCG GC S ATG TC narG-r GAG TTG TAC CAG TCR GCS GAY TCS G		Initial Denaturation	95ºC	15min	(Bru et al. 2007)
			5 cycles	95ºC 63ºC Decrease 1ºC per cycle	30s 30s	
			35 cycles	95ºC 63ºC	30s 30s	
nirS	cd3aF: GT(C/G) AAC GT(C/G) AAG GA(A/G) AC(C/G) GG R3cd: GA(C/G) TTC GG(A/G) TG(C/G) GTC TTG A	387bp	Initial Denaturation	95ºC	10min	(Throback et al. 2004)
			35 cycles	95ºC 56ºC 72ºC	30s 30s 30s	
nirK	nirK5R: GCC TCG ATC AGR TTR TGG nirK583FdegCF: TCA TGG TGC TGC CGC GYG ANG G	~430bp	Initial Denaturation	94ºC	2min	(Braker et al. 1998) (Santoro et al. 2006)
			35 cycles	94ºC 60ºC 72ºC	30s 60s 60s	
Clade I *nosZ*	NosZ1840f: CGC RAC GGC AAS AAG GTS MSS GT NosZ2090r: CAK RTG CAK SGC RTG GCA GAA	267bp	Initial Denaturation	95ºC	15min	(Henry et al. 2006)
			6 cycles	95ºC 65ºC Decrease 1ºC per cycle	15s 30s	
			35 cycles	95ºC 59ºC 72ºC 80ºC	30s 15s 30s 15s	

Gene	Primers	Size	Step	Temperature	Time	Reference
Clade II *nosZ*	nosZ-II-f: CTI GGI CCI YTK CAY AC nosZ-II-r: GCI GAR CAR AAI TCB GTR C	745bp	Initial Denaturation	95ºC	5min	(Jones et al. 2013)
			35cycles	95ºC 55ºC 72ºC 80ºC	30s 60s 45s 30s	

Table S3. Quality control parameters for qPCR assays.

Gene	Efficiency	R^2
Universal 16S rRNA	92.2%	0.999
Accumulibacter 16S rRNA	100.3%	0.997
narG	88.6%	0.992
nirS	96.2%	0.999
nirK	97.6%	0.997
Clade I *nosZ*	92.0%	0.998
Clade II *nosZ*	96.2%	0.993

Table S4. Relative abundance (%) of Accumulibacter 16S rRNA (Accumulibacter) and five denitrification genes (*narG*, *nirS*, *nirK* and clade I and II *nosZ*) normalized to universal bacterial 16S rRNA gene abundance quantified by qPCR assays.

	Operation days	**Accumulibacter**	***narG***	***nirS***	***nirK***	***nosZ* clade I**	***nosZ* clade II**
Total biomass	Day 458	6.1±0.5	1.1±0.2	7.5±1.0	0.6±0.2	3.0±0.4	0.5±0.2
	Day 472	8.0±2.1	1.9±0.0	13.3±5.7	1.1±0.3	4.4±0.7	1.2±0.1
	Day 486	7.5±1.9	2.9±1.1	11.5±4.6	0.5±0.3	4.8±3.3	0.7±0.2
	Day 507	8.7±2.4	1.7±0.7	12.5±3.2	0.5±0.0	2.2±0.8	2.2±1.7
	Day 521	8.2±2.1	1.2±0.1	7.5±6.6	0.9±0.5	3.1±1.1	2.1±1.2
	Day 535	5.5±1.3	1.4±0.8	5.2±0.9	0.7±0.0	2.6±0.4	1.8±0.3
	Day 549	4.2±1.2	2.3±0.5	9.7±1.7	0.9±0.5	4.2±1.7	2.1±0.4
	Day 563	8.1±2.5	2.9±0.6	8.7±1.5	0.7±0.0	1.9±0.4	2.4±1.3
Different fraction of biomass (μm)	Day 591 <150	5.2±0.6	3.1±0.8	9.0±4.8	0.5±0.2	3.6±0.1	1.4±0.4
	Day 591 150-350	6.6±4.7	2.9±0.7	26.4±17.5	0.8±0.6	7.1±5.4	2.6±1.3
	Day 591 350-600	16.3±7.1	3.5±2.0	18.2±8.2	2.9±0.6	4.7±0.4	11.8±8
	Day 591 >600	27.5±14.4	1.5±0.6	17.0±5.3	1.6±1.1	12.9±4.5	8.7±6.4
	Day 626 <150	7.1±0.6	3.9±0.1	6.4±0.9	0.6±0.1	4.7±0.1	1.9±0.3
	Day 626 150-350	14.9±2.5	4.5±0.8	21.5±1.7	0.9±0.1	4.0±0.4	1.9±0.0
	Day 626 350-600	21.5±4.5	4.5±2.4	16.6±8.3	1.5±0.7	5.8±0.4	6.7±0.2
	Day 626 >600	26.4±4.2	1.7±0.1	15.6±7.5	0.7±0.2	4.9±0.4	6.7±5.7

Supporting Information - Figures

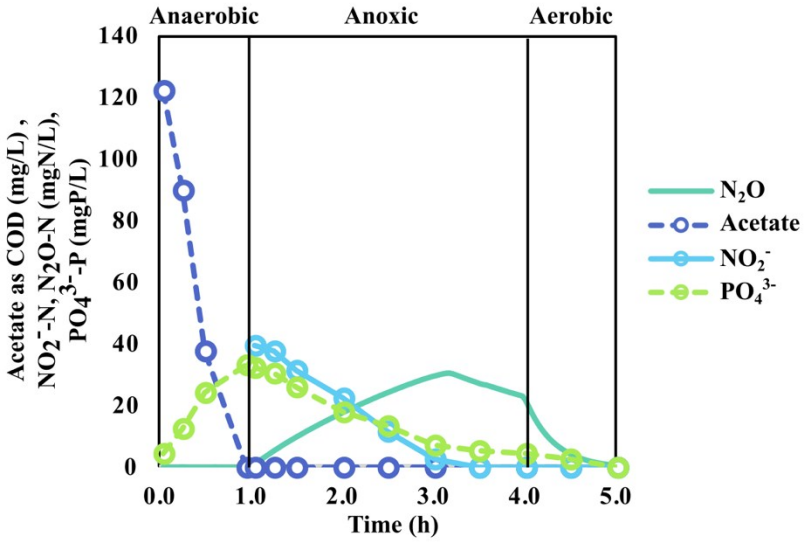
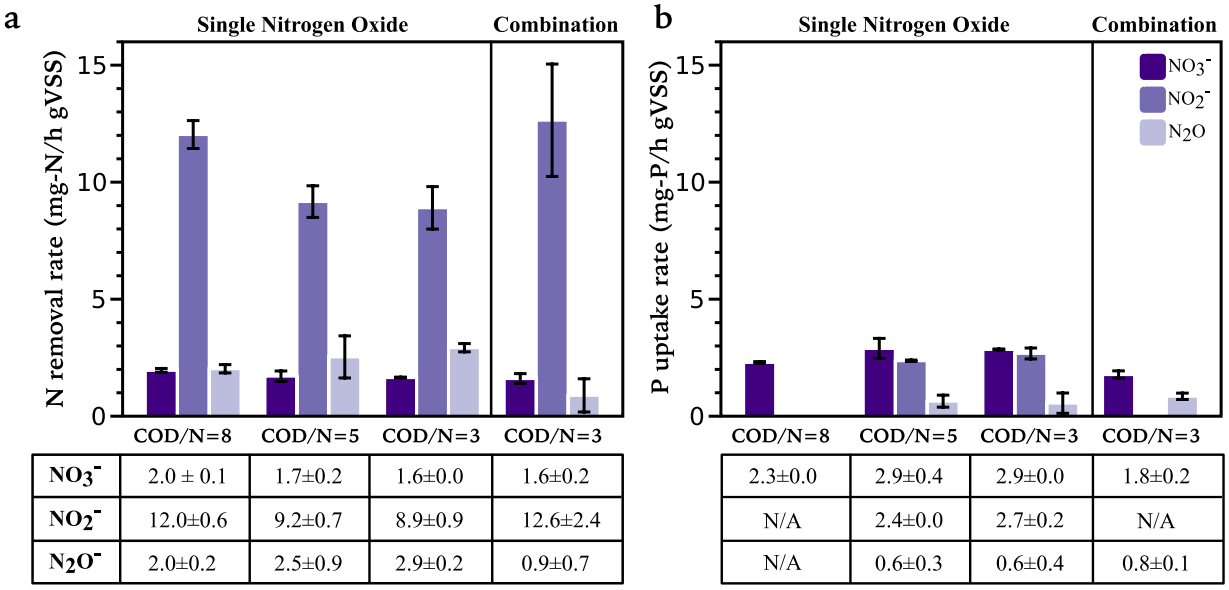


Figure S1. A typical SBR cycle profiling the transformation of key C, N and P components (acetate as COD, NO_2^-, dissolved N_2O and PO_4^{3-}).



NO_3^-	2.0 ± 0.1	1.7±0.2	1.6±0.0	1.6±0.2
NO_2^-	12.0±0.6	9.2±0.7	8.9±0.9	12.6±2.4
N_2O^-	2.0±0.2	2.5±0.9	2.9±0.2	0.9±0.7

2.3±0.0	2.9±0.4	2.9±0.0	1.8±0.2
N/A	2.4±0.0	2.7±0.2	N/A
N/A	0.6±0.3	0.6±0.4	0.8±0.1

Figure S2. a) N removal rate (mg-N/h-gVSS) and **b)** denitrifying P uptake rate (mg-P/h-gVSS) by DPAO-enriched biomass with different dosing schemes in coupled feeding mode (simultaneous addition of COD [acetate] and nitrogen oxide(s) with 40 mg-N/L of nitrogen oxides). x-axis labels refer to following conditions: single nitrogen oxide: COD/N=8 (320mg COD/L, C_a), COD/N=5 (200 mg COD/L, C_b), and COD/N=3 (120mg COD/L, C_c); combination: COD/N=3 (39 mg-N/L, combination of 13mg-N/L of each nitrogen oxide (NO_3^-, NO_2^- and N_2O), C_a_NO_3+NO_2+N_2O). The NO_2^- driven P uptake rate is unavailable (when COD/N = 8 and 5) due to the rapid reduction rate of NO_2^- and lack of electron acceptor after COD was completely consumed. The P uptake rate for N_2O when COD/N=3 is not shown due to the large standard deviation for the duplicate experiments.

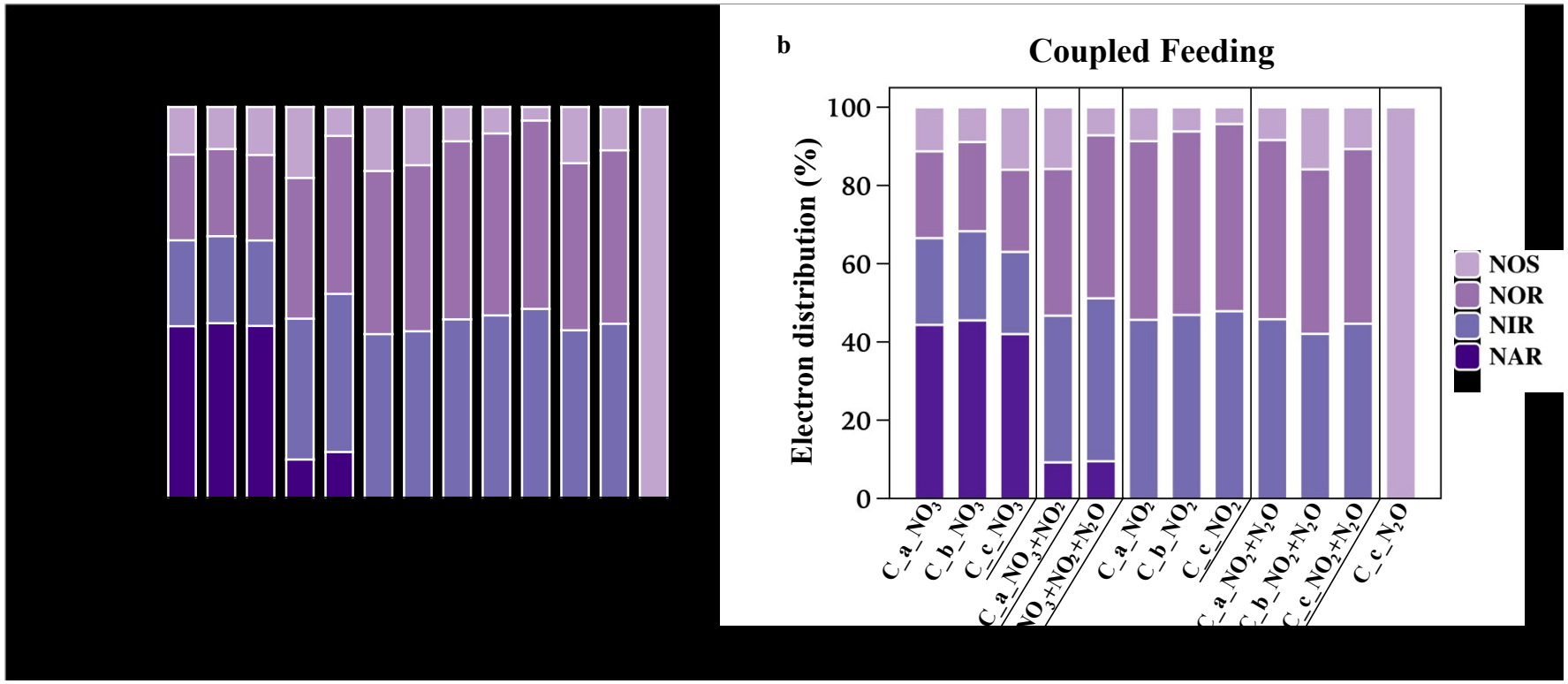


Figure S3. Electron distribution for different denitrification enzymes (NAR: nitrate reductase, NIR: nitrite reductase, NOR: nitric oxide reductase, NOS: nitrous oxide reductase) under (**a**) decoupled and (**b**) coupled feeding modes. Dosing schemes in x-axes in the figure correspond to batch test scenario in Table 1, and vary based on concentration and type of nitrogen oxides supplied as electron acceptor(s) (NO_3^-, NO_2^-, and/or N_2O). Dosing scheme: a,b,c_NOx: varying concentrations of NOx (10 mg-N/L, 20 mg-N/L and 40 mg-N/L); a_NO_3+NO_2+N_2O: 13 mg-N/L each of NO_3^-, NO_2^- and N_2O; a,b,c_NO_2+N_2O: different combinations of NO_2^- and N_2O; a_NO_3+NO_2: 20 mg-N/L NO_3^- and NO_2^-; D_d_NO_2: 60 mg N/L NO_2^-. Complete details of nitrogen oxide dosing concentrations can be found in Table 1.

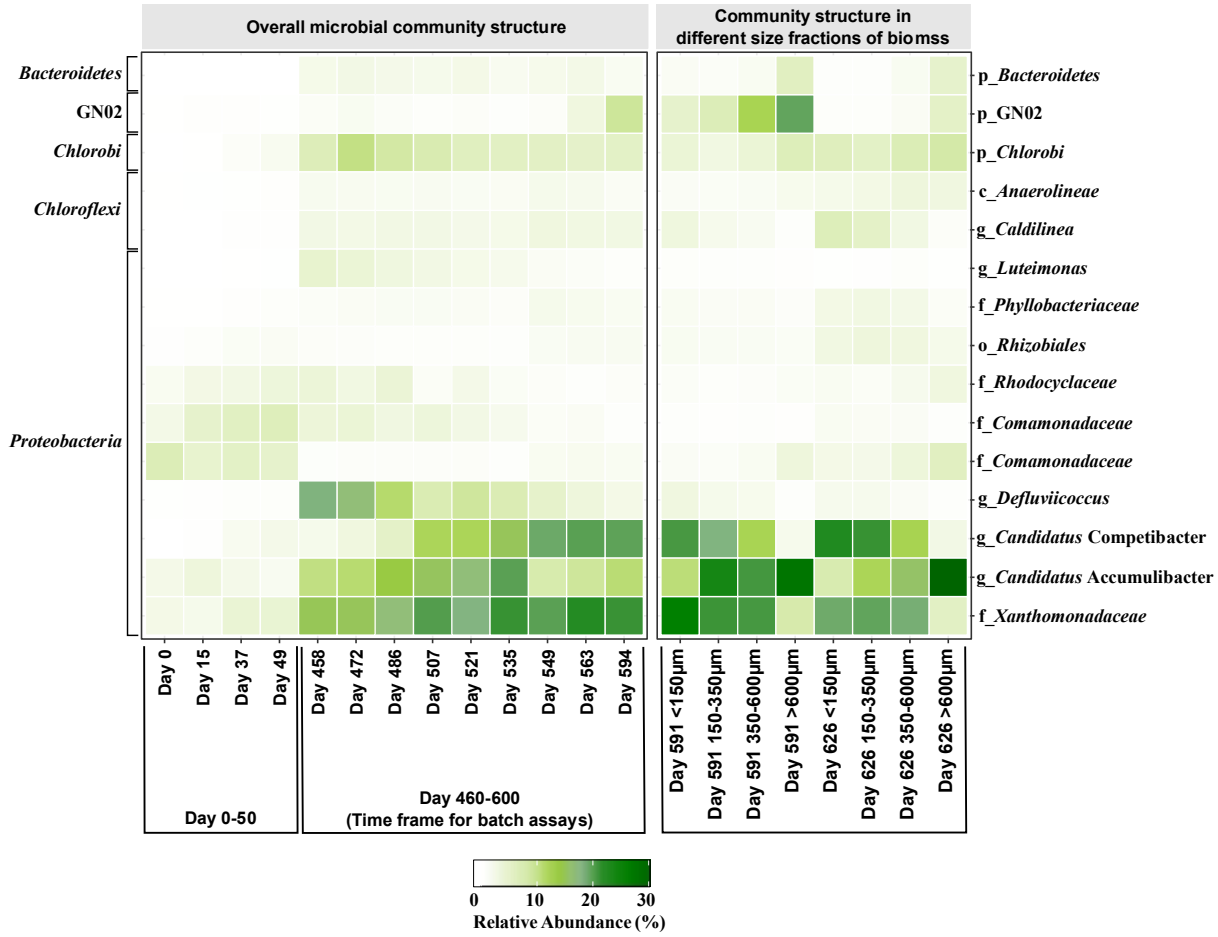


Figure S4. Relative abundance for the 15-most abundant microbial taxa at the genus level in the mother SBR via 16S rRNA gene amplicon sequencing. Left: Overall microbial community structure in biomass over time range when batch assays were conducted (day 460 to day 600) and within two months of SBR inoculation (day 0 to day 50); Right: Microbial community structure in different size fractions (<150 um, 150-350 um, 350-600 um, and >600 um) of biomass in two selected days (day 591 and 626). Taxonomy is shown at the phylum level (left label) and at the lowest level of taxonomic assignment (p: phylum, c: class, o: order, f: family, and g: genus; right label).

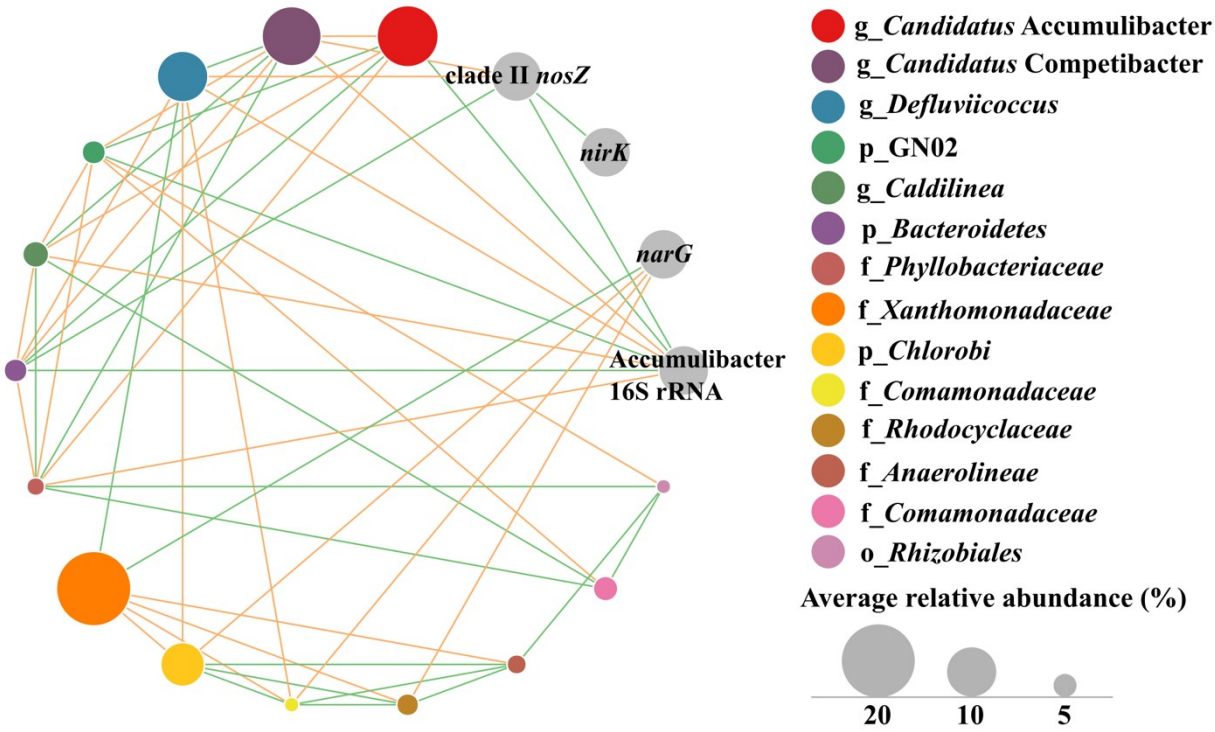


Figure S5. Correlation network between the top 15 most abundant bacterial taxa at the genus level detected in the nitrite-fed denitrifying EBPR bioreactor via 16S rRNA gene sequencing (colored circles) and the abundance of denitrification and Accumulibacter 16S rRNA (gray circles; *nirK, narG,* Accumulibacter 16S rRNA, and clade II *nosZ*) measured via qPCR. Spearman correlation analysis between the top 15 bacterial genera was conducted based on the relative abundance of these bacterial genera in different biomass size fractions. Taxa names are given at the lowest level of taxonomic assignment (p: phylum, c: class, o: order, f: family, and g: genus). *nirS* and clade I *nosZ* are not shown in the figure because they were not correlated with any bacterial taxon or denitrification genes. A positive correlation (Pearson correlation $\rho>0.8$, $p<0.05$) between nodes is indicated by a green edge, and negative correlation (Pearson correlation $\rho<-0.8$, $p<0.05$) is indicated by an orange edge. The size for each colored circle represents the average relative abundance (%) of the bacterial genus in reactor biomass between day 486 and day 625 (time frame when batch assays occurred).

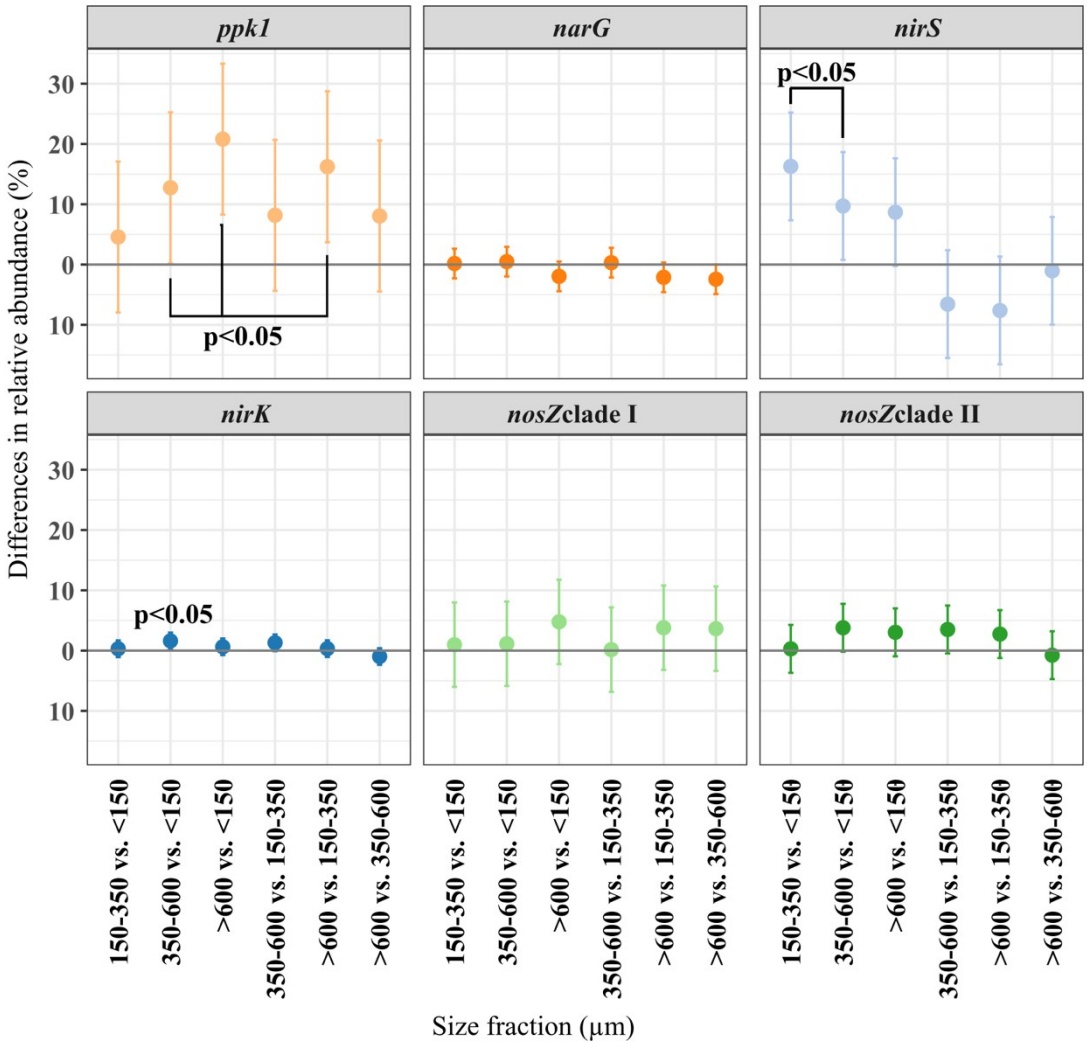


Figure S6. Average differences in functional gene relative abundance (%) between each biomass size fraction calculated based on post-hoc test. Error bar represents the 95% confidence interval. A positive number on the y-axis indicates higher abundance in the larger size fraction.

References

Braker, G., Fesefeldt, A. and Witzel, K.-P. (1998) Development of PCR Primer Systems for Amplification of Nitrite Reductase Genes (*nirK* and *nirS*) to detect denitrifying bacteria in environmental samples. Applied and Environmental Microbiology 64(10), 3769-3775.

Bru, D., Sarr, A. and Philippot, L. (2007) Relative Abundances of Proteobacterial Membrane-Bound and Periplasmic Nitrate Reductases in Selected Environments. Applied and Environmental Microbiology 73(18), 5971-5974.

Burgmann, H., Jenni, S., Vazquez, F. and Udert, K.M. (2011) Regime shift and microbial dynamics in a sequencing batch reactor for nitrification and anammox treatment of urine. Applied Environmental Microbiology 77(17), 5897-5907.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J. and Knight, R. (2010) QIIME allows analysis of high-throughput community sequencing data. Nature Publishing Group 7(5), 335-336.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G. and Knight, R. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME Journal 6(8), 1621-1624.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19), 2460-2461.

Edgar, R.C. (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Method 10(10), 996-998.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C. and Knight, R. (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27(16), 2194-2200.

Griffin, J.S. and Wells, G.F. (2017) Regional synchrony in full-scale activated sludge bioreactors due to deterministic microbial community assembly. ISME J 11(2), 500-511.

He, S., Gall, D.L. and McMahon, K.D. (2007) "*Candidatus* Accumulibacter" population structure in enhanced biological phosphorus removal sludges as revealed by polyphosphate kinase genes. Applied Environmental Microbiology 73(18), 5865-5874.

Henry, S., Bru, D., Stres, B., Hallet, S. and Philippot, L. (2006) Quantitative detection of the *nosZ* gene, encoding nitrous oxide reductase, and comparison of the abundances of 16S rRNA, *narG*, *nirK*, and *nosZ* genes in soils. Applied and Environmental Microbiology 72(8), 5181-5189.

Jones, C.M., Graf, D.R., Bru, D., Philippot, L. and Hallin, S. (2013) The unaccounted yet abundant nitrous oxide-reducing microbial community: a potential nitrous oxide sink. ISME J 7(2), 417-426.
Pruesse, E., Peplies, J. and Glockner, F.O. (2012) SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28(14), 1823-1829.

Santoro, A.E., Boehm, A.B. and Francis, C.A. (2006) Denitrifier community composition along a nitrate and salinity gradient in a coastal aquifer. Applied and Environmental Microbiology 72(3), 2102-2109.

Throback, I.N., Enwall, K., Jarvis, A. and Hallin, S. (2004) Reassessing PCR primers targeting *nirS*, *nirK* and *nosZ* genes for community surveys of denitrifying bacteria with DGGE. FEMS Microbiology Ecology 49(3), 401-417.