

Supporting Information

Whole community metagenomics in two different anammox configurations:
process performance and community structure

Ananda S. Bhattacharjee^{1¶}, Sha Wu^{1¶}, Christopher E Lawson², Mike SM Jetten³, Vikram Kapoor⁴, Jorge W. Santo Domingo⁵, Katherine D. McMahon^{2,6}, Daniel R. Noguera² and Ramesh Goel^{1}*

¹Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, UT 84112

²Department of Civil and Environmental Engineering, University of Wisconsin – Madison, 1415

Engineering Drive, Madison, WI 53706, USA

³Department of Microbiology, Radboud University, Nijmegen, The Netherlands.

⁴Department of Civil and Environmental Engineering, University of Texas at San Antonio, One UTSA
Circle, San Antonio, TX 78249

⁵U.S. Environmental Protection Agency, Office of Research and Development, National Risk
Management Research Laboratory, Cincinnati, Ohio 45268

⁶Department of Bacteriology, University of Wisconsin – Madison

* Corresponding author: Ph: 801-581-6110; ram.goel@utah.edu

[¶]Equal contribution by authors

Contents:

Table S1: Summarizes the MG-RAST analysis results for SGR and AGR metagenome datasets.

Table S2 (a): Lists the anammox metabolism genes in *Ca' Brocadia* sp 1

Table S2 (b): Lists the anammox metabolism genes in *Ca' Brocadia* sp 2

Figure S1: Schematic diagram of two-stages suspended (a) and attached (b) anammox reactors.

Figure S2: Nitrogen concentration ($\text{NH}_4^+\text{-N}$, $\text{NO}_2^-\text{-N}$, $\text{NO}_3^-\text{-N}$) in influent and effluent and the removal efficiency for partial nitrification in (a) suspended growth and (b) attached growth reactor.

Figure S3: Rarefaction curve of annotated species as a function of percentage of reads.

Figure S4: Taxonomic classification based on 16S rRNA gene sequences. Total of clustered 3,140 and 3,416 reads from SGR and AGR genome were attributed to 16S rRNA gene sequences, and assigned to 20 and 16 different phyla for SGR and AGR, respectively. Each slice represents the fraction of the total reads which was assigned to specific phylum. Three digits fractions in % for SGR and AGR were presented after each label in parentheses.

49 Table S1. Summarizes the MG-RAST analysis results for SGR and AGR datasets.

MG-RAST analysis	SGR	AGR
Total reads	8,247,168	8,865,646
Total reads uploaded	6,504,924	7,137,843
QC* passed reads (rRNA genes)	309,473 (4.8%)	387,548(5.4%)
QC* passed reads (predicted protein with known function)	3,589,563(55.2%)	4,102,652(57.5%)
QC* passed reads (predicted protein with unknown function)	319,664(4.9%)	463,456(6.5%)
QC* failed reads	2,286,224(35.1%)	2,184,187 (30.6%)

50 *Minimum complexity reads are placed under this node

51

52

53

54

55

56

57

58

59

60

61

62

63 Table S2(a). Lists the anammox metabolism genes in *Ca' Brocadia* sp 1

Locus tag	Start	End	Product
BROC1_16_6	7622	9295	hydrazine oxidoreductase <i>HzoA</i>
BROC1_16_1	1123	2430	hydrazine synthase A subunit
BROC1_16_0	2	1069	hydrazine synthase A subunit
BROC1_16_2	2844	4451	hydroxylamine oxidoreductase
BROC1_16_3	4608	5057	hydroxylamine oxidoreductase
BROC1_19_8	8832	9890	putative copper-type nitrite reductase
BROC1_19_7	8171	8866	putative copper-type nitrite reductase
BROC1_19_19	20642	23158	hydroxylamine oxidoreductase
BROC1_31_12	14619	15833	nitrite transporter protein
BROC1_36_18	19076	19441	hydrazine oxidoreductase <i>HzoA</i>
BROC1_51_10	7925	9601	hydrazine oxidoreductase <i>HzoA</i>
BROC1_52_15	20774	22075	hydrazine synthase A subunit
BROC1_52_16	22135	23202	hydrazine synthase A subunit
BROC1_60_8	9676	10575	nitrite/formate transporter
BROC1_60_7	8740	9660	nitrite/formate transporter
BROC1_85_18	14551	16797	hydroxylamine oxidoreductase
BROC1_85_17	13844	14605	hydroxylamine oxidoreductase
BROC1_98_6	6102	7349	hydroxylamine oxidoreductase
BROC1_98_20	23248	26652	hydroxylamine oxidoreductase
BROC1_98_6	6102	7349	hydroxylamine oxidoreductase
BROC1_98_6	6102	7349	hydroxylamine oxidoreductase
BROC1_98_20	20088	21698	hydroxylamine oxidoreductase
BROC1_98_7	7447	8298	hydroxylamine oxidoreductase
BROC1_153_15	19206	20336	hydrazine synthase C subunit
BROC1_153_16	20441	21280	hydrazine synthase B subunit
BROC1_154_40	35440	36456	hydrazine synthase B subunit
BROC1_154_41	36777	37061	hydrazine synthase C subunit
BROC1_164_58	55475	56839	ammonium transporter
BROC1_164_56	53441	54991	ammonium transporter
BROC1_164_10	10412	12088	hydroxylamine oxidoreductase
BROC1_181_33	38589	39854	hydroxylamine oxidoreductase
BROC1_182_18	10227	11141	nitrite/formate transporter
BROC1_191_89	86579	87955	hydroxylamine oxidoreductase
BROC1_191_87	85674	86180	hydroxylamine oxidoreductase

64

65

66

67 Table S2(b). Lists the anammox metabolism genes in *Ca' Brocadia* sp 2

Locus tag	Start	End	Product
BROC2_0_181	224779	226428	hydroxylamine oxidoreductase
BROC2_0_26	26549	28006	hydroxylamine oxidoreductase
BROC2_2_180	201636	203381	hydroxylamine oxidoreductase
BROC2_3_105	118501	120111	hydroxylamine oxidoreductase
BROC2_4_0	66	1196	hydrazine synthase C subunit
BROC2_10_0	1	609	hydrazine oxidoreductase <i>HzoA</i>
BROC2_10_50	58979	60514	hydroxylamine oxidoreductase
BROC2_11_0	3	1349	hydrazine synthase A subunit
BROC2_13_19	24974	26188	nitrite transporter protein
BROC2_18_66	58438	59568	hydrazine synthase C subunit
BROC2_19_23	24132	25556	ammonium transporter
BROC2_20_0	1	1395	hydrazine oxidoreductase <i>HzoA</i>
BROC2_23_0	1	654	hydrazine oxidoreductase <i>HzoA</i>
BROC2_32_4	5957	7303	hydrazine synthase A subunit
BROC2_32_0	1	1395	hydrazine oxidoreductase <i>HzoA</i>
BROC2_32_3	3947	5554	hydroxylamine oxidoreductase

68
69
70
71
72
73
74
75
76
77
78
79
80

Figure S1. Schematic diagram of two-stages suspended (a) and attached (b) anammox reactors.

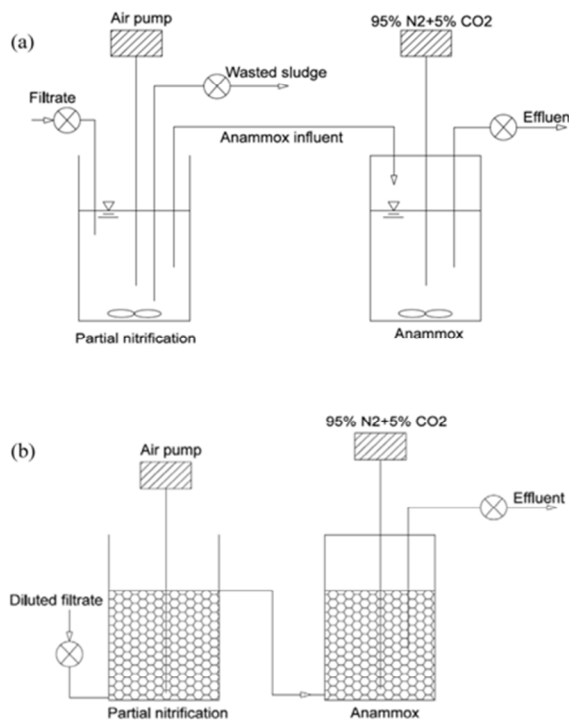
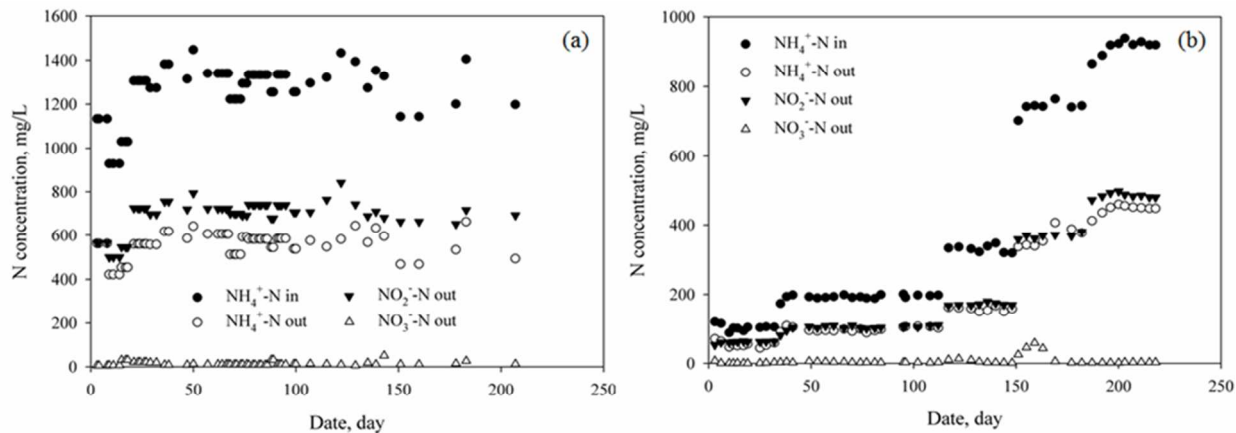


Figure S2. Nitrogen concentration ($\text{NH}_4^+\text{-N}$, $\text{NO}_2^-\text{-N}$, $\text{NO}_3^-\text{-N}$) in influent and effluent and the removal efficiency for partial nitritation in (a) suspended growth and (b) attached growth reactor.



89 Figure S3. Rarefaction curve of annotated species as a function of percentage of reads.

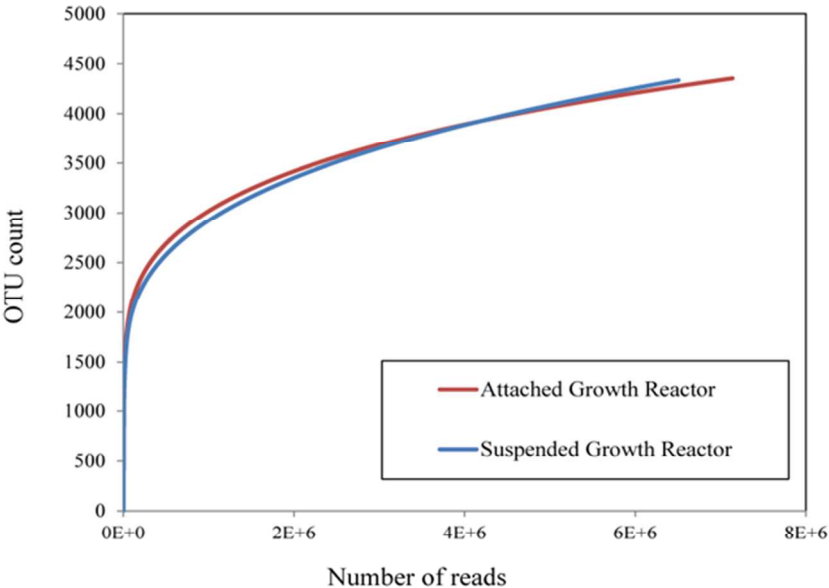


Figure S4. Taxonomic classification based on 16S rRNA gene sequences. Total of clustered 3,140 and 3,416 reads from SGR and AGR genome were attributed to 16S rRNA gene sequences, and assigned to 20 and 16 different phyla for SGR and AGR, respectively. Each slice represents the fraction of the total reads that was assigned to specific phylum. Three digits fractions in % for SGR and AGR were presented after each label in parentheses.

