Long solids retention times and attached growth phase favor prevalence of comammox bacteria in nitrogen removal systems.

Irmarie Cotto[1], Zihan Dai[2], Linxuan Huo[1], Christopher L. Anderson[1], Katherine J. Vilardi[1], Umer Ijaz[2], Wendell Khunjar[3], Christopher Wilson[4], Haydee De Clippeleir[5], Kevin Gilmore[6], Erika Bailey[7], Ameet J. Pinto[1*].

1 Department of Civil and Environmental Engineering, Northeastern University.
2 School of Engineering, University of Glasgow.
3 Hazen and Sawyer, Inc.
4 Hampton Roads Sanitation District.
5 DC Water.
6 Department of Civil and Environmental Engineering, Bucknell University.
7 City of Raleigh Public Utilities.

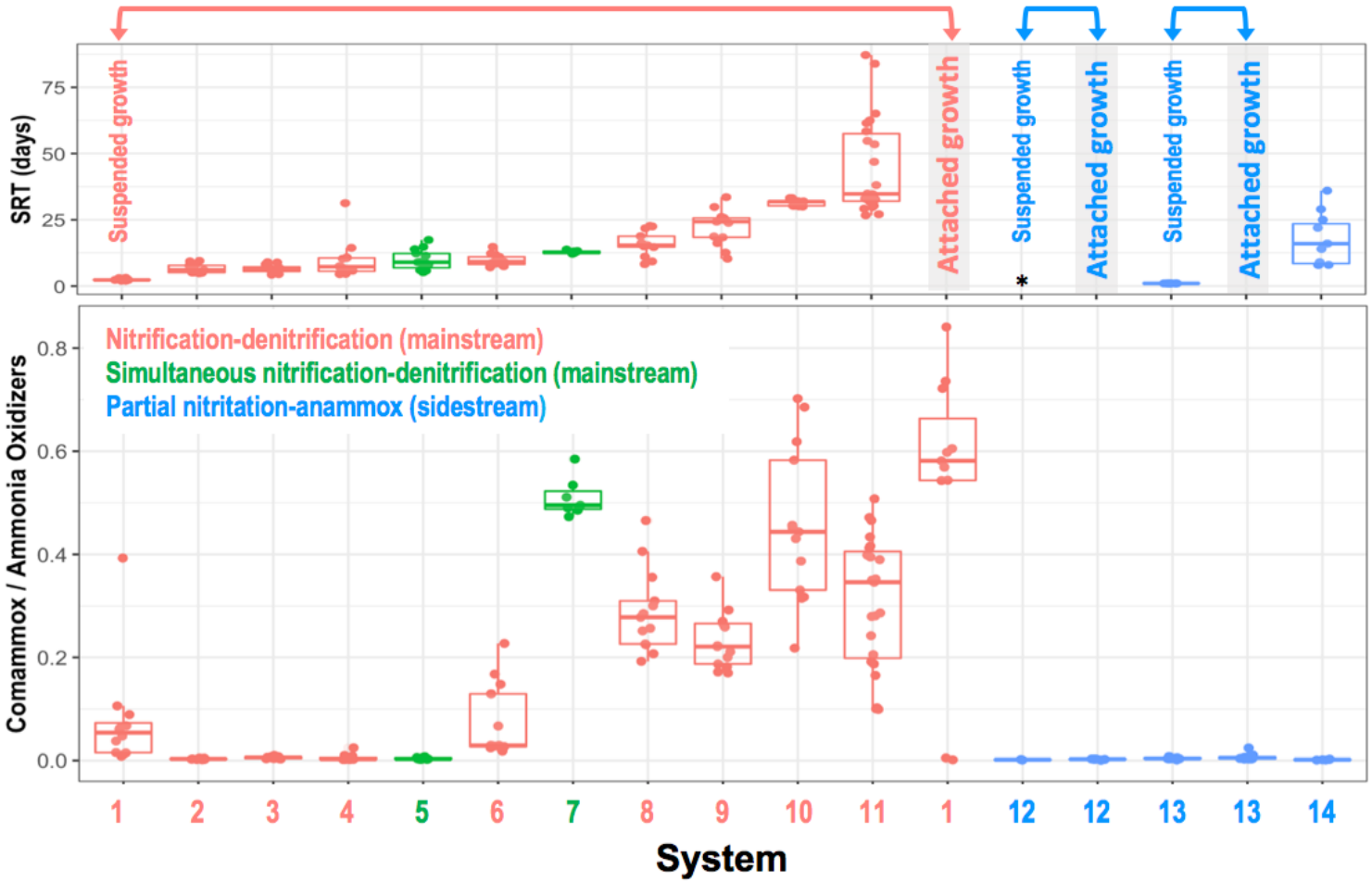
*corresponding author: a.pinto@northeastern.edu

Keywords: comammox bacteria, nitrification, solids retention time, qPCR, metagenomics

Highlights

- Clade A comammox bacteria were detected in wastewater nitrogen removal systems.
- New qPCR assay targeting the *amoB* gene of clade A comammox bacteria was developed.
- Comammox bacteria are prevalent in mainstream conventional and simultaneous nitrification-denitrification systems with long solids retention times (>10 days).
- Comammox bacteria were not detected in sidestream partial nitrification-anammox systems included in this study.

GRAPHICAL ABSTRACT



ABSTRACT

The discovery of the complete ammonia oxidizing (comammox) bacteria overturns the traditional two-organism nitrification paradigm which largely underpins the design and operation of nitrogen removal during wastewater treatment. Quantifying the abundance, diversity, and activity of comammox bacteria in wastewater treatment systems is important for ensuring a clear understanding of the nitrogen biotransformations responsible for ammonia removal. To this end, we conducted a yearlong survey of 14 full-scale nitrogen removal systems including mainstream conventional and simultaneous nitrification-denitrification and side-stream partial nitrification-anammox systems with varying process configurations. Metagenomics and genome-resolved metagenomics identified comammox bacteria in mainstream conventional and simultaneous nitrification-denitrification systems, with no evidence for their presence in side-stream partial nitrification-anammox systems. Further, comammox bacterial diversity was restricted to clade A and these clade A comammox bacteria were detected in systems with long solids retention times (>10 days) and/or in the attached growth phase. Using a newly designed qPCR assay targeting the *amoB* gene of clade A comammox bacteria in combination with quantitation of other canonical nitrifiers, we show that long solids retention time is the key process parameter associated with the prevalence and abundance of comammox bacteria. The increase in comammox bacterial abundance was not associated with concomitant decrease in the abundance of canonical nitrifiers; however, systems with comammox bacteria showed significantly better and temporally stable ammonia removal compared to systems where they were not detected. Finally, in contrast to recent studies, we do not find any significant association of comammox bacterial prevalence and abundance with dissolved oxygen concentrations in this study.

1.0 Introduction

Nitrification, the oxidation of ammonia to nitrate via nitrite, coupled with denitrification where nitrate is reduced to dinitrogen gas are key processes in the removal of nitrogen from wastewater (Klotz and Stein 2008). Traditionally, nitrification was considered as a two-step process driven by two distinct nitrifying guilds, i.e., ammonia oxidation by the aerobic ammonia oxidizing bacteria (AOB) (Kowalchuk and Stephen 2001) or archaea (AOA) (Stahl et al. 2012) followed by nitrite oxidation by the aerobic nitrite oxidizing bacteria (NOB) ((Daims et al. 2016). Ammonia oxidizers within the *Betaproteobacteriales* genera *Nitrosomonas* and *Nitrosospira* (Siripong and Rittmann 2007, Wu et al. 2019) and NOB within the genus *Nitrospira* (Juretschko et al. 1998, Wu et al. 2019) and more recently *Nitrotoga* (Saunders et al. 2016) are thought to be dominant nitrifiers in wastewater treatment systems. However, the discovery of complete ammonia oxidizing (i.e., comammox) bacteria (Daims et al. 2015, Pinto et al. 2015, van Kessel et al. 2015) has added new complexity to nitrogen biotransformation in wastewater systems. For instance, comammox bacteria can completely oxidize ammonia to nitrate and may compete with canonical AOB and NOB, as well as anaerobic ammonia oxidizing (anammox) bacteria in partial nitritation-anammox (PNA) systems, also known as deammonification systems. The potential competition for ammonia and possibly nitrite among comammox bacteria and other nitrifiers could have implications for process design and operation not only in wastewater treatment but also across other ecosystems.

Comammox bacteria have been detected in geothermal springs (Daims et al. 2015), aquaculture systems (van Kessel et al. 2015), drinking water (Pinto et al. 2015, Tatari et al. 2017, Wang et al. 2017), rapid gravity sand filters treating groundwater (Fowler et al. 2018, Palomo et al. 2016), soils (Hu and He 2017, Orellana et al. 2018), as well as a range of wastewater treatment bioreactors (Annavajhala et al. 2018, Camejo et al. 2017, Chao et al. 2016, Fan et al. 2017, Gonzalez-Martinez et al. 2016, Pjevac et al. 2017, Roots et al. 2019, Spasov et al. 2019, Wang et al. 2018, Xia et al. 2018). The key to evaluating the impact of comammox bacteria on wastewater process operations is to understand the impact of key process variables on whether comammox prevalence and abundance and in turn how this impacts overall activity and function of the engineered system. This can then help delineate laboratory-, pilot-, or even full-scale experiments to probe competitive dynamics between comammox bacteria and other nitrifiers in scenarios that are relevant from a process operations perspective. To address this issue, the current study presents a systematic year-

long evaluation of nitrifying populations, including comammox bacteria, in full-scale wastewater treatment plants to provide a baseline of process configurations, operations, and environmental conditions under which comammox bacteria might be important.

All detected comammox bacteria belong to genus *Nitrospira* (lineage II) and exhibit close phylogenetic relatedness to canonical *Nitrospira*-NOB. Recent work suggests that the capacity to oxidize ammonia via the ammonia monoxygenase (AMO) and hydroxylamine dehydrogenase (HAO) enzymes may have been acquired by comammox *Nitrospira* via horizontal gene transfer (Palomo et al. 2018). Nonetheless, the close phylogenetic affiliation of comammox bacteria with canonical NOB represents a major challenge with their detection and quantitation. Specifically, the 16S rRNA gene and subunits A (*nxrA*) and B (*nxrB*) of the nitrite oxidoreductase (NXR) genes cannot distinguish between comammox bacteria from canonical NOB within the genus *Nitrospira.* One alternative to identify and obtain relative abundance of comammox bacteria within a complex nitrifying consortium involves the use of shotgun DNA sequencing (i.e., metagenomics). This has been employed by several studies (Annavajhala et al. 2018, Camejo et al. 2017, Chao et al. 2016, Fan et al. 2017, Gonzalez-Martinez et al. 2016, Pinto et al. 2015, Roots et al. 2019, Spasov et al. 2019, Xia et al. 2018) to demonstrate that comammox bacterial presence in wastewater systems is primarily dominated by clade A comammox bacteria closely related to *Ca* Nitrospira nitrosa. While metagenomics provides a snapshot overview, it is not ideally suited for high-throughput profiling of large number of samples particularly for microbial groups that are typically low-to-medium abundance (e.g., nitrifiers) due to sequencing cost and sequencing depth issues. To this end, the ammonia monooxygenase (amo) gene sequences provide a convenient approach for detection and quantitation. Specifically, both the subunit's A (*amoA*) and B (*amoB*) of the ammonia monooxygenase genes form distinct clusters from other known AOB and AOA. This sequence divergence has been used to develop primer sets targeting the *amoA* gene of clade A and clade B comammox bacteria separately (Pjevac et al. 2017) and together (Bartelme et al. 2017, Fowler et al. 2018, Wang et al. 2018). A consistent challenge with these primers is the formation of unspecific products and our experience in the study suggests that they are also unable to capture comammox bacteria detected via metagenomics (see results section). Alternatively, (Beach and Noguera 2019) proposed the use of species specific primers that target comammox bacterial *amoA* gene depending on the process of interest. For instance, they proposed that since *Ca* Nitrospira

125 nitrosa are dominant in low energy wastewater treatment systems, utilizing primers that capture
126 *Ca* Nitrospira nitrosa would be ideal. While this circumvents the challenge of unspecific product
127 formation, a species-specific approach eliminates the possibility of detecting other closely related
128 comammox bacteria. In this study, we used a metagenomic approach to recover *amoA* and *amoB*
129 genes from several full-scale nitrogen removal systems and use them in combination with
130 previously published gene sequences to design and validate clade level primers for the *amoB* gene
131 of clade A comammox bacteria; this is to our knowledge the primary comammox clade of
132 relevance of wastewater treatment systems.
133
134 The overall objectives of this study were (1) to identify nitrogen removal process configurations
135 for wastewater treatment where comammox bacteria are likely to be relevant, (2) develop a qPCR
136 assay for quantitation of comammox bacteria in complex nitrifying communities, and (3) perform
137 a temporal survey across a range of process configurations to determine the influence of
138 environmental and variable process operation conditions on the abundance of nitrifying
139 populations inclusive of comammox bacteria. This was accomplished through year-long
140 quantitative tracking of nitrifying populations in fourteen nitrogen removal systems with varying
141 process configuration using genome resolved metagenomics and qPCR combined with appropriate
142 statistical analyses.
143
144 **2.0 Materials and Methods**
145 **2.1 Sampling sites, sample processing, and data collection.**
146 Samples were collected on a monthly basis from June 2017 to June 2018 at fourteen nitrogen
147 removal systems which included seven full-scale suspended growth nitrification-denitrification
148 (ND) systems, one full-scale integrated fixed film activated sludge (IFAS) ND system, two full-
149 scale simultaneous ND (SND) systems, one full-scale suspended growth side stream PNA system,
150 and two full-scale IFAS PNA systems. Table 1 provides an overview of the nitrogen removal
151 systems sampled in this study. Sample collection and processing were conducted according to the
152 MIDAS field guide (McIlroy et al. 2015). Specifically, samples were collected from the nitrifying
153 bioreactors by operational personnel at each of the wastewater utilities and 100 ml of sample was
154 shipped overnight in coolers with icepacks to Northeastern University (NU). Immediately upon
155 arrival to NU, the samples were homogenized using the Hei-TORQUE Value 400 (Heidolph, Cat.

156 No. 036093070) in a 30-mL glass/Teflon tissue grinder (DWK Life Sciences Wheaton, Cat. No.
157 357984) for 1 minute (2nd gear, speed 9, 10 times from top to bottom of the glass tissue grinder)
158 following the MIDAS protocol, and four 1 mL homogenized aliquots per sample were transferred
159 to Lysing Matrix E tubes (MP Biomedical, Cat. No. 6914100). Samples were centrifuged at
160 10,000g for 5 minutes and the supernatant was discarded and biomass pellet was stored at -80°C
161 until DNA extraction. Further, ~20 mg of media attached biomass samples from the IFAS system
162 were also stored in Lysing Matrix E tubes at -80°C. In addition, the utilities provided data on pH,
163 temperature, ammonia, nitrite, nitrate, chemical oxygen demand (COD), biological oxygen
164 demand (BOD), total suspended solids (TSS), volatile suspended solids (VSS), mixed liquor TSS
165 (MLSS) and mixed liquor VSS (MLVSS) from the influent, nitrification reactor, and effluent for
166 each sampling time point, as well as hydraulic retention time (HRT), solids retention time (SRT)
167 and dissolved oxygen (DO) data from the nitrification reactors.
168

169 **Table 1**: Overview of sampling locations, sampling scheme used in this manuscript, process type and sub-type,
170 operational scale and sampling time-frame for this study.

Site code	Process type	Process sub-type	Treatment stream	Sampling Start (mm/yy)	Sampling End (mm/yy)
NEU	ND[1]	Four/five stage	Mainstream	06/17	06/18
DUR	PNA[2]	Anitamox	Sidestream	06/17	06/18
YAN	SND[3]	ORP control[a]	Mainstream	11/17	05/18
GRE	ND[1]	SBR[b]	Mainstream	06/17	06/18
DCWM	ND[1]	Five stage	Mainstream	06/17	06/18
KIN	ND[1]	MBR/MLE[c]	Mainstream	06/17	06/18
WIL	ND[1]	Five stage	Mainstream	06/17	06/18
JAMM	ND[1]	IFAS[d]	Mainstream	06/17	06/18
JAMS	PNA[2]	Anitamox	Sidestream	06/17	06/18
YORM	ND[1]	Conventional	Mainstream	06/17	06/18
YORS	PNA[2]	DEMON	Sidestream	06/17	06/18
BOA	ND[1]	High MLSS	Mainstream	06/17	06/18
NAN	SND[3]	Five stage	Mainstream	06/17	06/18
ARM	ND[1]	Five stage	Mainstream	06/17	06/18

171 [1]Nitrification-denitrification, [2]Partial nitritation-anammox, [3]Simultaneous nitrification-denitrification, [a]Oxidative-
172 Reduction Potential Control, [b]Sequencing batch reactor, [c]Membrane bioreactor/Modified Ludzack-Ettinger, [d]Integrated
173 Fixed-film Activated Sludge
174

175

176

177 **2.2 DNA extraction.**

178 Samples were subject to DNA extraction using the DNAeasy PowerSoil kit (Qiagen, Inc., Cat.
179 No.12888) with automated extraction instrument QIAcube (Qiagen, Inc., Cat. No. 9002160)
180 following manufacturer instructions with some modifications in the sample preparation step.
181 Briefly, the lysing buffer from the PowerBead Tubes provided in the DNAeasy PowerSoil kit was
182 transferred to the lysing matrix tubes containing the samples and 60 µL of Solution C1 from the
183 kit was added to each sample. To complete cell lysis, bead-beating was performed using a
184 FastPrep-24 instrument (MP Bio, Cat. No. 116005500) four times for 40 seconds each (McIlroy
185 et al. 2015). Between each 40 second bead beating interval, samples were kept on ice for 2 min to
186 prevent excess heating. Samples were centrifuged at 10,000g for 30 seconds and the supernatant
187 was further purified using the QIAcube Protocol Sheet for the DNeasy PowerSoil Kit. Ten samples
188 and two blanks (only reagents) were extracted for each DNA extraction run. Extracted DNA was
189 quantified using Qubit instrument with Qubit dsDNA Broad Range Assay (ThermoFisher
190 Scientific, Cat. No. Q32850) and a subset of samples were randomly selected for analysis using
191 1% agarose gel electrophoresis to visualize DNA shearing. Extracted DNA concentrations ranged
192 from 2 to 254 ng/µL. DNA extracts were stored at -80°C until further analysis.

193

194 **2.3 Metagenomic sequencing and data analyses.**

195 One sample from each system was selected for metagenomic sequencing based on qPCR estimates
196 of AOB and *Nitrospira* (see qPCR details in section 2.4). Specifically, samples with high
197 *Nitrospira*:AOB ratios were selected for metagenomic sequencing under the assumption that these
198 samples were likely to consist of comammox *Nitrospira* bacteria. For systems that included
199 distinct suspended and attached growth phase (e.g., IFAS system from JAMM and PNA systems
200 from JAMS and DUR), one sample each for the attached and suspended phase were included in
201 the sequencing run resulting in a total of 18 metagenomes. Extracted DNA was shipped frozen to
202 the Roy J. Carver Biotechnology Center at University of Illinois Urbana-Champaign Sequencing
203 Core for sequencing. The DNA extracts were subject to PCR-free library preparation using Hyper
204 Library construction kit from Kapa Biosystems and subsequently sequenced on 300 cycle run
205 (2x150nt reads) on two lanes of Illumina HiSeq 4000. This resulted in a total of 1.35 billion paired-
206 end reads. Raw data for these 18 metagenomes was deposited in NCBI with bioproject number
207 PRJNA552823. The reads were subject to adaptor removal and quality trimming using

208 Trimmomatic (Bolger et al. 2014) which included adaptor removal, clipping of the first and last
209 three bases, trimming of sequences where four-bp sliding window quality threshold was below
210 Q20, and discarding of all trimmed sequences less than 75 bp in length. This resulted in a total of
211 1.03 billion paired end reads which were analyzed using a few different approaches (see below).
212
213 **2.3.1 Reference genomes based annotation.**
214 A total of 53 publicly available Genomes/genome assemblies for nitrifying organisms were
215 downloaded from NCBI (RefSeq version 86) (Table S1). This included genome assemblies for 12
216 AOA, five anammox bacteria, 17 AOB, 10 comammox bacteria, and nine NOB. Reads from all
217 samples were competitively mapped to the reference genomes using bwa (Li and Durbin 2010)
218 followed by extraction of properly paired mapped reads (samtools view -f2) using SAMtools (Li
219 et al. 2009). The extracted reads were then again competitively aligned against reference genomes
220 using BLAST (Altschul et al. 1990) with a criterion of 90% sequence identity and 90% read
221 coverage. The reads per kilobase million (RPKM) metric was used as a measure of the reference
222 genome relative abundance in each sample and was estimated by dividing the number of reads
223 aligning to reference genome at aforementioned criteria with the number of kilobases in each
224 reference genome and the millions of reads per sample.
225
226 **2.3.2 SSU rRNA based community characterization.**
227 The paired end metagenomic reads were assembled using MATAM (Pericard et al. 2018) with the
228 SILVA SSU rRNA (Release 132) database (Quast et al. 2013) as the reference for assembly of the
229 16S rRNA genes with a minimum 16S rRNA gene length threshold of 500 bp. The assembled 16S
230 rRNA genes were classified against the RDP database (Cole et al. 2014) using Naïve Bayesian
231 classifier approach (Wang et al. 2007). The relative abundance of each 16S rRNA gene was
232 estimated by dividing the number of reads mapping to each MATAM assembled 16S rRNA gene
233 with the total number of reads mapping to all assembled 16S rRNA genes per sample.
234
235 **2.3.3 Gene-centric de novo assembly.**
236 The paired end reads from each sample assembled into contigs using metaSpades (Nurk et al.
237 2017); where two sets of reads were available from attached and suspended phase from same
238 bioreactor (i.e., JAMM, JAMS, DUR) or from two parallel operated bioreactors (GRE), these were

239 co-assembled. MetaSpades assembly/co-assembly was carried out with kmers 21, 33, 55, 77, 99,
240 and 127. The assembled contigs were filtered to remove all contigs smaller than 500 bp and subject
241 to gene calling using prodigal (Hyatt et al. 2010) and annotation against the KEGG database
242 (Kanehisa et al. 2016) using DIAMOND (Buchfink et al. 2015). Genes identified as encoding for
243 *amoA* (KO:K10944), *amoB* (KO:K10945), and *nxrA* (KO:K00370) were extracted for further
244 analyses. An analysis of genes assembled using *de novo* assembly approach for the complete
245 metagenome revealed that several of the identified genes were highly fragmented. To circumvent
246 this limitation, gene-centric *de novo* assembly was used. Specifically, paired end reads were
247 mapped to a curated amino acid database of *amoA*, *amoB*, and nxrA genes using DIAMOND
248 (Buchfink et al. 2015). The mapped reads were then assembled into contigs metaSpades (Nurk et
249 al. 2017), followed by gene calling using prodigal (Hyatt et al. 2010) and annotation against the
250 KEGG database (Kanehisa et al. 2016). The *amoA*, *amoB*, and *nxrA* genes identified using the
251 complete and gene-targeted *de novo* assembly approach were combined, de-duplicated, and further
252 curated by evaluating phylogenetic placement of assembled genes. Specifically, reference
253 alignments for *amoA*, *amoB*, and *nxrA* genes were created by extracting corresponding genes from
254 the reference assemblies and aligning with MUSCLE (Edgar 2004) followed by construction of a
255 reference tree for each gene using RAxML (Stamatakis 2014). The *amoA*, *amoB*, and *nxrA* genes
256 were placed on the gene-specific reference tree using pplacer (Matsen et al. 2010). Annotated
257 genes that did not conform to known phylogeny of *amoA* and *amoB* for AOB, AOA, and
258 comammox bacteria and *nxrA* genes for NOB and comammox bacteria were discarded. Finally,
259 the relative abundance of the curated genes in each sample was estimated using the RPKM metric.
260 Specifically, the sum of reads mapping to contig containing gene of interest per sample were
261 divided by number of kilobases in for each contig and the millions of reads per sample.
262

263 **2.3.4 Recovery, phylogenomic placement, and annotation of Nitrospira metagenome**
264 **assembled genomes.**

265 Assembled contigs from nine metagenomes corresponding to seven nitrogen removal systems
266 where comammox bacteria were detected (i.e., DCWM, GRE1/GRE2, JAMM/JAMMSM, KIN,
267 NEU, WIL, and YAN) were pooled. Reads from each metagenome were mapped to all contigs
268 with bowtie2 (default parameters, version 2.1.0) (Langmead and Salzberg 2012) prior to binning
269 with MetaBAT2 (-m 2000, version 2.12.1) (Kang et al. 2015). The completion and redundancy of

270 the resulting bins were estimated with CheckM (lineage_wf, version 1.0.12) (Parks et al. 2015).
271 Reads mapping to bins ≥ 50% complete were extracted and used to re-assemble the bin with
272 Unicycler (default parameters, version 0.4.7) (Bankevich et al. 2012, Li et al. 2009, Wick et al.
273 2017). The quality and taxonomy of re-assembled bins were evaluated with CheckM (Parks et al.
274 2015) and the Genome Taxonomy Database Toolkit (GTDB-Tk 0.2.2, database release r86 v3)
275 (Parks et al. 2018). Bins greater than 70% complete and classified as Nitrospira by GTDB-Tk were
276 retained for manual refinement with Anvi'o (Eren et al. 2015). Following refinement, the quality
277 and taxonomy of the bins were re-assessed with CheckM (Parks et al. 2015) and the Genome
278 Taxonomy Database Toolkit (Parks et al. 2018), selecting metagenome assembled genomes
279 (MAGs) with completeness and redundancy estimates of ≥ 70% and ≤ 10%, respectively. Open
280 reading frames (ORFs) were predicted using Prodigal (Hyatt et al. 2010) v2.6.3 for all *Nitrospira*
281 MAGs recovered from this study (n=10). KEGG orthologies (KO) were assigned to predicted
282 ORF's in these 10 genomes against 10,108 HMM models of prokaryota in KEGG(Kanehisa et al.
283 2016) database v90.1 using kofamscan (Aramaki et al. 2019). To investigate the phylogeny of
284 these 10 *Nitrospira* MAGs recovered from this study, 32 previously publicly available N*itrospira*
285 genomes were downloaded from Genbank (Table S2) (Poghosyan et al. 2019) and used as
286 reference genomes for phylogenetic tree reconstruction. Phylogenomic tree reconstruction was
287 conducted by Anvi'o (Eren et al. 2015) v5.5. ORFs were predicted for aforementioned 32 reference
288 genomes and 10 MAGs from this study using Prodigal (version 2.6.3) (Hyatt et al. 2010) and
289 then searched against a collection of HMM models summarized by Campbell et al. (Campbell et
290 al. 2011) using hmmscan (version 3.2.1) (Eddy 2011) including 48 ribosomal proteins. Only
291 genomes containing more than genes encoding for 40 of the 48 ribosomal proteins were included
292 in downstream phylogenomic analyses. Alignments for each gene were conducted using muscle
293 (Edgar 2004), alignments were concatenated, and finally phylogenomic tree was constructed using
294 FastTree (version 2.1.7) (Price et al. 2010).
295

296 **2.4 Quantitative PCR for total bacteria and nitrifying populations and design of primers**
297 **targeting *amoB* gene of clade A comammox bacteria.**

298 PCR thermocycling and reaction mix conditions of previously published primer sets for qPCR-
299 based quantification of 16S rRNA gene of total bacteria (Caporaso et al. 2011), AOB (Hermansson
300 and Lindgren 2001) and *Nitrospira* (Graham et al. 2007) and ammonia monoxygenase subunit A

301 (*amoA*) gene for AOB (Rotthauwe et al. 1997) are shown in Table 2. Further, qPCR assays
302 targeting the *amoA* gene of comammox bacteria were conducted using previously published primer
303 sets (Fowler et al. 2018, Pjevac et al. 2017) and additional comammox *amoA* primer sets designed
304 a part of this study (Table 2). However, these assays either resulted in unspecific product formation
305 and non-detection of comammox bacteria in samples where metagenomic analyses indicated
306 presence of comammox bacteria. As a result, new primer sets targeting the *amoB* gene of clade A
307 comammox bacteria (only clade A comammox bacteria were detected via metagenomic analyses
308 in this study) were designed. To do this, the *amoB* genes assembled from the metagenomic data
309 were combined with *amoB* gene sequences from previous studies (Daims et al. 2015, Palomo et
310 al. 2016, Pinto et al. 2015, van Kessel et al. 2015) and two primer sets were designed (Table 2).
311 These new set of primers were tested with samples that were positive and negative for comammox
312 bacteria, along with DNA extracts from *Ca* Nitrospira inopinata as positive control involving
313 variation in annealing temperature, primer concentration, and template concentration.
314
315 The qPCR assays were performed on a QuantStudio 3 Real-Time PCR System (ThermoFisher
316 Scientific, Cat. No. A28567) in 20 µL reaction volume including: 10 µL Luna Universal qPCR
317 Master Mix (New England Biolabs, Inc., Cat. No. NC1276266), primers listed in Table 2, 5 µL of
318 10 times diluted DNA template and the required volume of DNAse/RNAse-Free water (Fisher
319 Scientific, Cat. No. 10977015) to reach 20 µL reaction. Reactions were prepared by the epMotion
320 M5073 liquid handling system (Eppendorf, Cat. No. 5073000205D) in triplicate. The cycling
321 conditions were as follows: initial denaturing at 95°C for 1 min followed by 40 cycles of denaturing
322 at 95°C for 15 s, annealing temperatures and times listed in Table 2, and extension at 72°C for 1
323 min. Melting curve analyses was performed at 95°C for 15 s, 60°C for 1 min, and 95°C for 15 s. A
324 negative control (NTC) and a standard curve ranging from 10^3-10^9 copies of 16S rRNA gene of
325 *Nitrosomonas europaea* for total bacteria assay and 10^2 – 10^8 copies of 16S rRNA genes for
326 *Nitrosomonas europaea* and *Ca* Nitrospira inopinata for the AOB and *Nitrospira* assays,
327 respectively and 10^2 – 10^8 copies for *amoA* of *Nitrosomonas europaea* and *Ca* Nitrospira inopinata
328 for the AOB and comammox assays, respectively and *amoB* gene of *Ca* Nitrospira inopinata for
329 the comammox assays were included in the qPCR analysis.
330
331

332 **2.5 Statistical analyses.**

333 Statistically significant differences in the abundance of nitrifying organisms between nitrogen
334 removal systems was evaluated using the non-parametric Kruskal-Wallis or Wilcoxon rank sum
335 test in R as appropriate (RCoreTeam 2014). Correlations between any two variables were
336 determined using Spearman rank correlations and BIOENV analysis within the R package "vegan"
337 to identify process and environmental variables that demonstrate significant correlation with
338 changes in nitrifier population abundances. All statistical tests and figure generation (Wickham
339 2009) were performed in R (RCoreTeam 2014).

340 **Table 2**: A summary of qPCR primers used and tested as part of this study (*equimolar proportions of forward and/or reverse primers if multiple primers are used).

Primer set	Assay	Gene	Forward primer	Reverse primer	Product size (bp)	Annealing Temp (ºC)/Time (sec)	Reference
F515 - R806	Total bacteria	16S rRNA	GTGCCAGCMGCCGCGGTAA	GGACTACHVGGGTWTCTAAT	291	50/15	(Caporaso et al. 2011)
CTO189FA/B/C* - RT1R	AOB	16S rRNA	GGAGRAAAGCAGGGGATCG GGAGGAAAGTAGGGGATCG	CGTCCTCTCAGACCARCTACTG	116	57/30	(Hermansson and Lindgren 2001)
amoA1F - 2R	AOB	*amoA*	GGGG TTTCTACTGGTGGT	CCCCTCKGSAAAGCCTTCTTC	491	54/30	(Rotthauwe et al. 1997)
Nspra675F - 746R	Nitrospira	16S rRNA	GCGGTGAAATGCGTAGAKATCG	TCAGCGTCAGRWAYGTTCCAGAG	93	58/30	(Graham et al. 2007)
comaA-D 244F - 659R	Comammox (clade A)	*amoA*	TAYAAYTGGGTSAAYTA	ARATCATSGTGCTRTG	415	52/45	(Pjevac et al. 2017)
comaB-D 244F - 659R	Comammox (clade B)	*amoA*	TAYTTCTGGACRTTYTA	ARATCCARACDGTGTG	415	52/45	(Pjevac et al. 2017)
comaA 244F* - 659R*	Comammox (clade A)	*amoA*	TACAACTGGGTGAACTA TATAACTGGGTGAACTA TACAATTGGGTGAACTA TACAACTGGGTCAACTA TACAACTGGGTCAATTA TATAACTGGGTCAATTA	AGATCATGGTGCTATG AAATCATGGTGCTATG AGATCATGGTGCTGTG AAATCATGGTGCTGTG AGATCATCGTGCTGTG AAATCATCGTGCTGTG	415	52/45	(Pjevac et al. 2017)
comaB 244F*- 659R*	Comammox (clade B)	*amoA*	TAYTTCTGGACGTTCTA TAYTTCTGGACATTCTA TACTTCTGGACTTTCTA TAYTTCTGGACGTTTTA TAYTTCTGGACATTTTA TACTTCTGGACCTTCTA	ARATCCAGACGGTGTG ARATCCAAACGGTGTG ARATCCAGACAGTGTG ARATCCAAACAGTGTG AGATCCAGACTGTGTG AGATCCAAACAGTGTG	415	52/45	(Pjevac et al. 2017)
Ntspa-amoA 162F - 359R	Comammox	*amoA*	GGATTTCTGGNTSGATTGGA	WAGTTNGACCACCASTACCA	197	52/45	(Fowler et al. 2018)
comaA - F/R	Comammox (clade A)	*amoA*	CARTGGTGGCCBATCGT	TTNGACCACCACCACCA	170	52/45	This study
comaB - F/R	Comammox (clade B)	*amoA*	GGNGACTGGGAYTTCTGG	GCCCACCARTACCARGC	190	52/45	This study
Cmx_amoB_1	Comammox (clade A)	*amoB*	TGACSATGGATAABGAGG	TCCGGATCGTGRAGAATGTC TCTGGATCGTGRAGAATGTC	114	52/45	This study
Cmx_amoB_2	Comammox (clade A)	*amoB*	TGGTAYGAYACNGAATGGG	CCCGTGATRTCCATCCA	337	52/45	This study

341

3.0 Results and Discussion

3.1 Overview of nitrogen removal systems included in this study.

As part of this study, we sampled fourteen full-scale nitrogen removal systems of varying process configurations. Specifically, we sampled nine nitrification-denitrification (ND), two simultaneous nitrification-denitrification (SND) mainstream systems, and three partial nitrification-anammox (PNA) sidestream systems with diverse process subtypes (Table 1). The process sub-types ranged from multi-stage suspended growth systems with secondary clarification to sequencing batch reactors (SBRs) ones to membrane bioreactors (KIN) to systems with significant attached growth components (e.g., DUR, JAMM, JAMS). The typical process parameters for each of these systems is shown in Figure 1. All nitrogen removal systems included in this study performed normally during this study with no significant and sustained process upset.

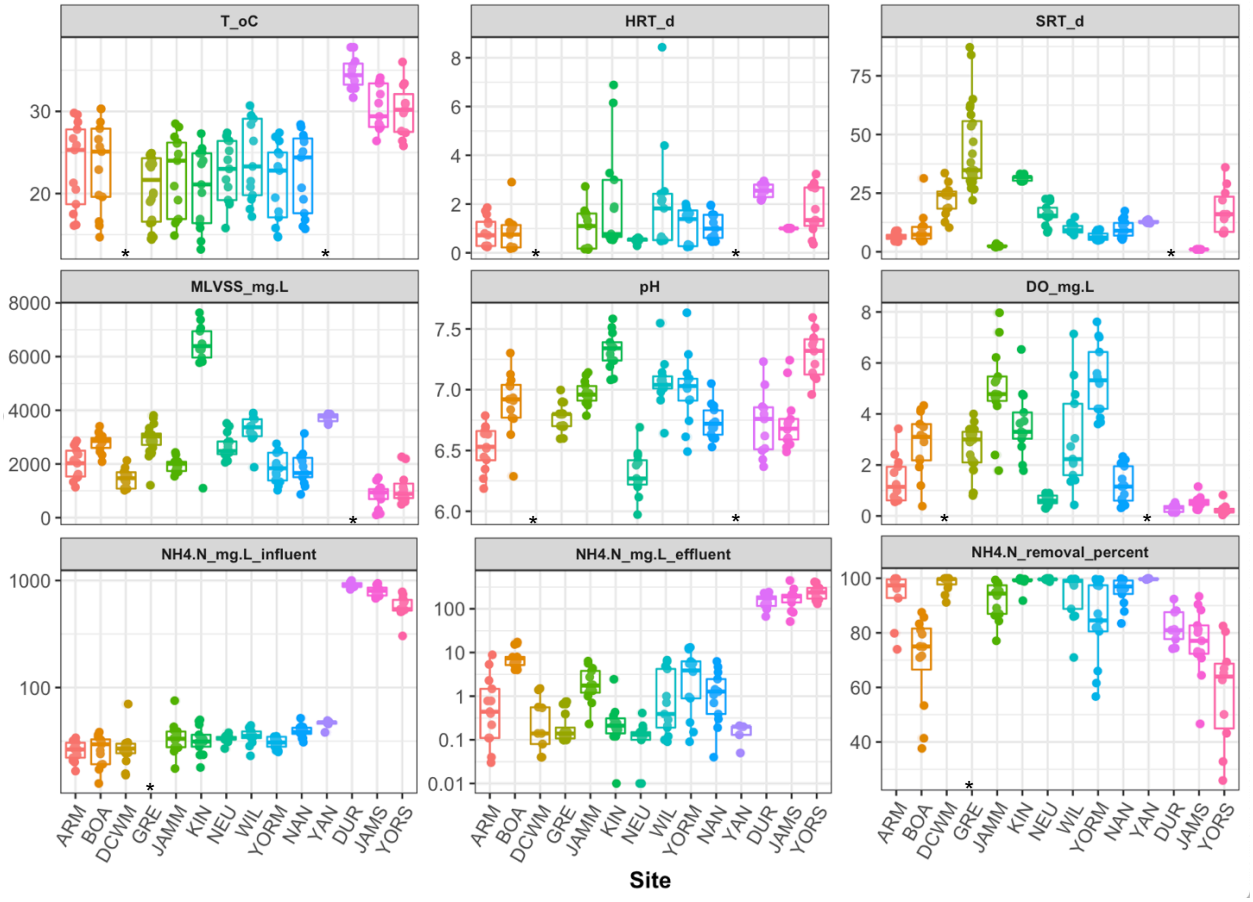


Figure 1: Overview of key process parameters and performance metrics monitored as part of this study for 14 nitrogen removal systems. Column indicated with "*" indicates lack of relevant process data for the respective nitrogen removal system.

353 **3.2 Identification of nitrifying populations using reference genome mapping and 16S rRNA**
354 **gene assembly.**
355 Mapping of reads to reference genomes indicated that *Nitrospira*-like bacteria were dominant
356 members of the nitrifying communities in all ND and SND systems followed by *Nitrosomonas*-
357 like bacteria, with no or very low detection of AOA (Figure 2A). Nearly all detected AOB
358 belonged to the genus *Nitrosomonas* with low levels of detection of AOB within the genera
359 *Nitrosococcus* and *Nitrosospira* in two systems (i.e., ARM and BOA). Metagenomic reads mapped
360 to comammox references genomes within the genus *Nitrospira* for samples from six ND and one
361 SND system, with abundances higher than canonical *Nitrospira*-NOB for three ND systems (i.e.,
362 GRE, KIN, and NEU) and one SND system (i.e., YAN). Genera containing other canonical NOB,
363 i.e. *Nitrotoga* and *Nitrobacter*, were detected in two ND and one SND system with *Nitrotoga* being
364 the primary NOB in BOA. Mapping of metagenomic reads from PNA systems to reference
365 genomes primarily resulted in the detection of anammox bacteria (genus: *Brocadia*) and AOB
366 (genus: *Nitrosomonas*), with no detectable presence of *Nitrospira*. It is important to note that while
367 metagenomic reference genome based approach may be useful for detection of organisms, its
368 quantitative value is limited. This limitation likely emerges from the reliance of a reference
369 database of genomes and the stringent criteria used for read mapping (i.e., 90% sequence identity
370 between metagenomic read over 90% of the read length). Thus, it is possible that the significant
371 over representation of *Nitrospira* in the metagenomic reference genome-based approach emerges
372 from a high level of genomic similarity between *Nitrospira* in the samples and in the reference
373 database and under representation of AOB emerges from a low level of genomic similarity between
374 AOB in the samples and in the reference database. This is the most likely scenario because
375 genomic similarity between organisms within the same genus can be as low as 75% (Jain et al.
376 2018).
377
378 In contrast, a subset of genes (particularly ribosomal genes) may be far more reliable for relative
379 quantitation of microbial abundance within the larger community. Thus, we used MATAM to
380 identify reads originating from the 16S rRNA gene within each metagenomic sample and
381 subsequently assemble them and finally determine the relative abundance of each gene in the
382 sample microbial community. The reference genome mapping results were largely consistent with
383 classification of MATAM-assembled 16S rRNA genes (Figure 2B) with *Nitrospira*- and

384 *Nitrosomonas*-like bacteria dominant in ND and SND systems and *Brocadia*- and *Nitrosomonas*-
385 like bacteria being dominant in PNA systems. Three of the 14 systems included in this study
386 consisted of distinct suspended and attached growth phase, i.e., ND IFAS system (i.e., JAMM)
387 and two side-stream PNA systems (DUR and JAMS). Relative abundances of *Nitrospira*- and
388 *Nitrosomonas*-like bacteria were similar between the suspended and attached growth phase for
389 JAMM while *Brocadia*-like bacteria were enriched in the attached growth phase for DUR and
390 JAMS. The key differences between reference genome- and 16S rRNA gene analyses were for the
391 PNA systems and one ND system (i.e., BOA). Specifically, *Nitrospira*-like bacteria were also
392 detected in all PNA system included in this study using 16S rRNA gene analyses compared to
393 reference mapping of metagenomics reads and while *Nitrolancea*-like bacteria were the primary
394 nitrifying bacteria in BOA, 16S rRNA gene analyses suggested high abundance of *Brocadia*-like
395 bacteria in KIN as well. The discrepancies between reference genome based analyses and 16S
396 rRNA gene assembly based analyses may result from the limitation of the prior (as stated above),
397 inefficient assembly of 16S rRNA genes due to highly conserved regions (Miller et al. 2011,
398 Pericard et al. 2018), or a combination of both.

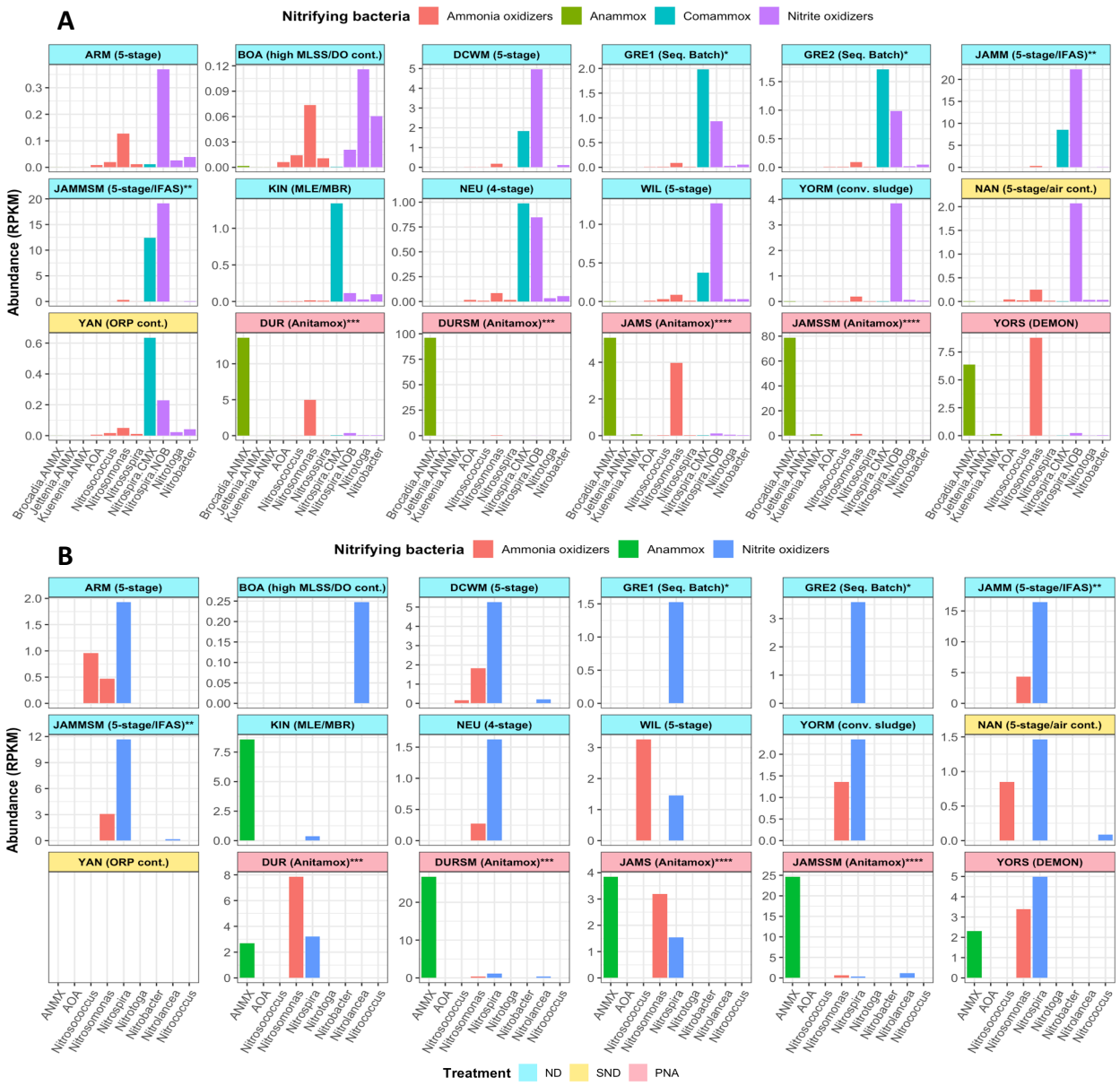


Figure 2: Relative abundance of nitrifying organisms in nitrogen removal systems based on (A) mapping of reads to reference genomes and (B) classification of MATAM assembled 16S rRNA genes. Bar colors indicate type of nitrifying organism detected (see color legend below each figure), while facet color indicates process type (i.e., red: PNA, Blue: ND, and Yellow: SND). GRE1/GRE2* panels present abundance of nitrifying populations from two different reactors at the same treatment plant. JAMM/JAMMSM**, DUR/DURSM***, and JAMS/JAMSSM**** panels present abundance of nitrifying populations in the suspended phase and attached growth phase of the system, respectively.

403 **3.3 Functional gene based metagenomic identification of ammonia and nitrite oxidizing**
404 **bacteria.**
405 Mapping of reads to curated *amoA*, *amoB*, and *nxrA* genes, followed by gene-centric *de novo*
406 assembly was used to detect and estimate the relative abundance of AOB, comammox bacteria,
407 and NOB in the metagenomic datasets from the 14 systems. Subsequently, all assembled and
408 annotated genes were placed on gene specific phylogenetic trees to eliminate potentially
409 misannotated *amoA*, *amoB*, and *nxrA* genes. This analysis revealed that nearly all annotated AOB
410 belonged to *Betaproteobacteriales* (Figure 3A) and contained representative sequences from ND,
411 SND, and PNA systems, while *amoA* and *amoB* sequences from comammox bacteria were
412 detected in DCWM, GRE, JAMM/JAMMSM, KIN, NEU, WIL, and YAN (Figure 3B). These
413 results are qualitatively consistent (i.e., presence/absence of comammox bacteria) with those
414 obtained by reference genome based analyses of comammox bacteria. The relative abundance of
415 comammox bacteria was higher than AOB in JAMM/JAMMSM, while both were equally
416 abundant two ND systems (i.e., GRE and NEU) and one SND system (i.e., YAN). In contrast,
417 comammox were less abundant than AOB in DCWM and WIL; both of these are ND systems.
418 Interestingly, no AOB were detected in KIN (an MBR/MLE system) with comammox being the
419 only ammonia oxidizer. The results for KIN were consistent with that of reference genome based
420 and 16S rRNA gene based analyses, where no AOB were detected. Further, the similarity in
421 relative abundance of *amoA*, *amoB*, and *nxrA* genes indicates that all detected *Nitrospira* bacteria
422 in this system are likely to be comammox bacteria. The non-detection of other AOB (e.g.,
423 *Nitrosospira*, *Nitrosococcus*) and NOB (e.g., *Nitrotoga*) that were detected by reference genome
424 based or 16S rRNA gene assembly based analyses could be due to challenges with *de novo*
425 assembly genes from low abundance microorganisms (i.e., insufficient sequencing depth and thus
426 low coverage).
427
428 A consistent aspect of systems with prevalent comammox populations was their high total solids
429 retention time (SRT). For instance, KIN is an MBR system with an estimated SRT of ~30 days,
430 DCWM is a ND system with a total SRT of ~30 days, and GRE is an SBR system with an SRT >
431 30 day. While JAMM/JAMMSM (i.e., James River) has an SRT of 2.5 days for the suspended
432 phase, this is an IFAS system with a significant attached growth component. This suggests that
433 irrespective of the process configuration type (ND, SND) and/or mode of operation (continuous,

434 SBR, attached growth), the key similarity among comammox prevalent systems is their high SRT
435 or presence of attached growth component (i.e., implicitly high SRT). This is consistent with
436 thermodynamic and metabolic pathway modeling that suggests that comammox bacteria have a
437 very low specific growth rates compared to AOB and NOB (Costa et al. 2006). For instance, recent
438 literature suggests that comammox maximum specific growth rates are 3-10, 5-11, and 2 times
439 lower than that of Nitrospira-NOB, AOB, and AOA (Kits et al. 2017, Lawson and Lücker 2018).
440 While this explains the prevalence of comammox in long SRT system, it does not explain the
441 absence of AOA in these systems. While AOA are highly prevalent in other ecosystems (i.e., soils,
442 drinking water, marine environments) (Prosser and Nicol 2008), they are rarely detected in
443 wastewater systems (Chao et al. 2016, Pjevac et al. 2017) including those in this project. It is
444 plausible that the absence of AOA and presence of comammox may be explained by the difference
445 in their affinity for ammonia, i.e., AOA with K_m (NH_3) ranging from 0.36 to 4.4 μM compared
446 with reported *Ca* Nitrospira inopinata K_m (NH_3) of 0.049 μM (Kits et al. 2017, Lawson and Lücker
447 2018).

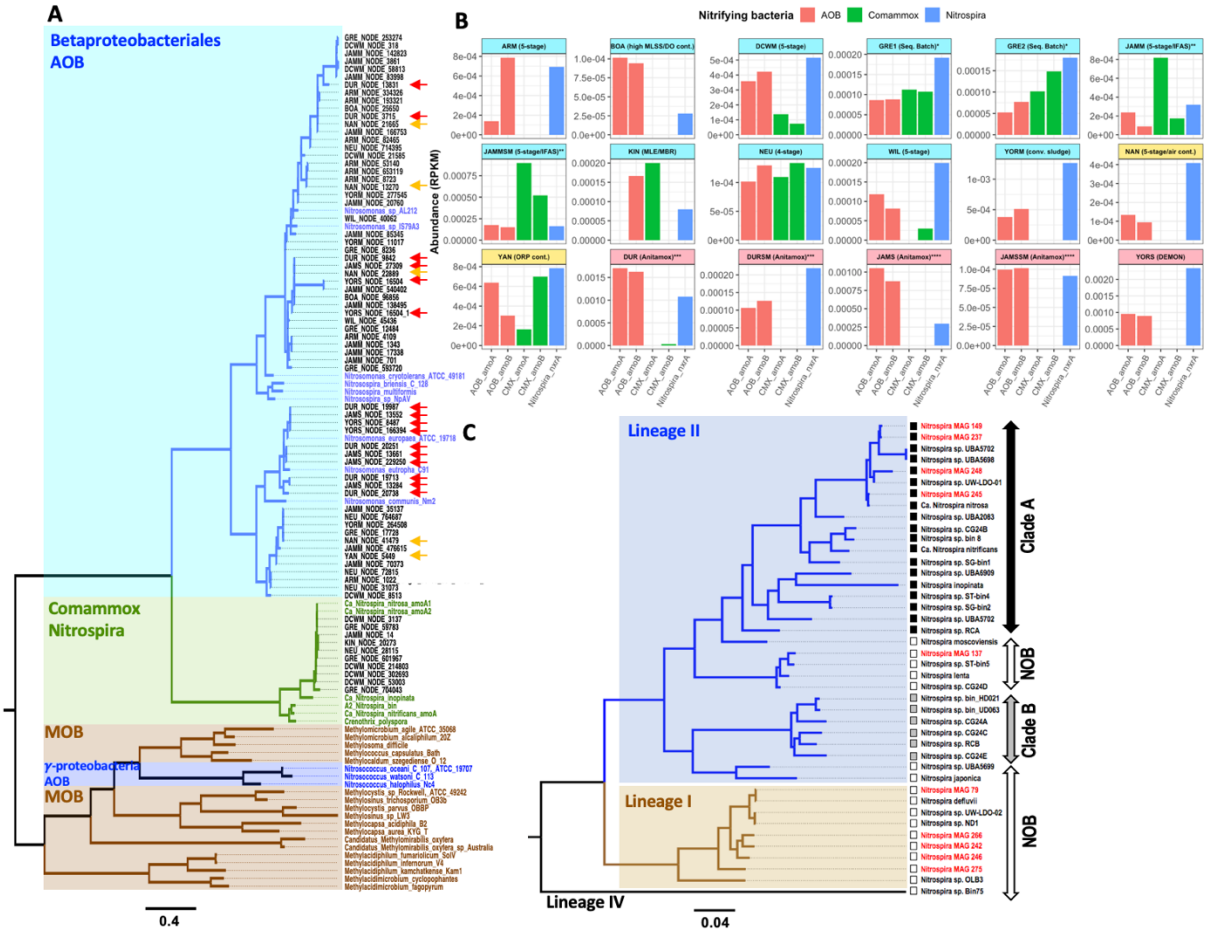


Figure 3: (A) Maximum likelihood phylogenetic tree of *amoA* genes recovered from gene centric *de novo* assembly along with reference *amoA* and *pmoA* gene sequences. Clades are colored by functional grouping and labels for reference sequences within each cluster are colored according to the functional group, while amoA sequences recovered from this study are shown in black labels. Red and orange arrows indicate amoA gene sequences recovered from PNA and SND systems respectively, while those recovered from ND systems are not annotated. amoA gene sequences from YAN and KIN are not shown in the phylogenetic tree due to recovery of short sequences not suitable for tree construction. (B) Abundance (RPKM) of amoA, amoB gene sequences of AOB and comammox bacteria, and nxrA gene sequences of Nitrospira are shown. (C) Phylogenetic placement of Nitrospira MAGs (red label) with 32 reference genomes (black label). Clade A (black squares), clade B (grey squares) comammox bacteria and canonical NOB (open squares).

All comammox based *amoA* and *amoB* gene sequences detected using metagenomic sequencing closely clustered with *Ca* Nitrospira nitrosa and distinct from other clade A bacteria (e.g., *Ca* Nitrospira nitrificans, *Ca* Nitrospira inopinata) (Figure 3B). Further, all Nitrospira MAG's recovered from the genome binning process were associated with either lineage 1 or 2 *Nitrospira* and included six canonical NOB and four clade A comammox. Genome statistics for the recovered MAGs are shown in Table 3. Consistent with previously described clade A comammox bacteria,

456 all comammox MAGs contained some or all genes associated with both ammonia and nitrite
457 oxidation, except for *Nitrospira* MAG 248 which is likely due to lower completeness (i.e., 70%).
458 Similarly, these comammox MAG's contained genes involved in urea uptake and hydrolysis - a
459 feature shared with the other recovered canonical *Nitrospira* NOB MAGs. Further, similar to other
460 clade A comammox bacteria, the recovered comammox MAG's lacked ammonium transporter
461 (Amt) or any of formate dehydrogenase genes which are reported present in clade B comammox
462 as well as canonical NOB within lineage II *Nitrospira* (Poghosyan et al. 2019). Finally, similar to
463 previously reported clade A comammox bacteria, the comammox MAG's recovered in this study
464 also indicated presence of genes associated with group 3b [NiFe] hydrogenase.
465
466 **Table 3**: Statistics of the Nitrospira MAG's assembled in this study and their functional grouping.

***Nitrospira* MAGs**	**79**	**137**	**149**	**237**	**242**	**245**	**246**	**248**	**266**	**275**
Nitrospira lineage	1	2	2	2	1	2	1	2	1	1
Completeness (%)	94.94	87.92	75.23	93.3	89.09	82.17	84.94	70.99	71.75	97.67
Redundancy (%)	1.36	9.15	1.14	3.86	2.36	4.6	4.09	2.12	1.75	1.36
Genome size (Mbp)	3.49	4.07	3.08	4.54	4.02	3.8	3.78	2.59	1.88	3.91
GC content	59.14	58.05	54.87	54.83	58.41	54.64	58.36	54.86	58.42	58.79
Number of contigs	136	410	229	48	114	203	98	412	179	54
N50 of contigs	40833	25623	21078	171941	177531	34975	85503	8266	19532	794701
genes	3361	4115	3141	4498	3886	3779	3694	2693	1842	3753
Protein coding genes	3322	4062	3113	4455	3845	3743	3653	2668	1821	3708
rRNA	0	1	1	3	0	0	0	0	0	0
tRNA	38	51	26	40	41	36	41	25	21	44
Functional group	NOB	NOB	CMX	CMX	NOB	CMX	NOB	CMX	NOB	NOB

467
468 **3.4 Development of qPCR assay for quantitative detection of comammox bacteria.**
469 We used previously developed primer sets (Fowler et al. 2018, Pjevac et al. 2017) to quantify the
470 abundance of comammox bacteria in samples collected as part of the study (Table 2). Pjevac et al.
471 (2017) provided two different primer sets targeting the *amoA* gene both clade A and clade B
472 comammox bacteria; one primer set was the degenerate primer set while the other contained an
473 equimolar proportion of six forward and six reverse primers. In contrast, Fowler et al. (2017)
474 developed a single primer set targeting the *amoA* gene of both clade A and clade B comammox

475 bacteria. While both primer sets showed excellent PCR efficiency with DNA extracts from pure
476 culture of *Ca* Nitrospira inopinata, efforts to PCR amplify the *amoA* gene from samples included
477 in this study demonstrated either non-detection of comammox bacteria in the six systems with
478 metagenomic evidence of their presence or unspecific product formation (Figure 4A, B). This was
479 also consistent for *amoA* specific primers designed in this study (Table 2). The same issue was
480 also highlighted by Beach and Noguera, (2019), who suggested the use of species primers for
481 comammox detection based on the system of interest (i.e., wastewater, drinking water, etc). The
482 comammox *amoA* primers were also tested against the comammox *amoA* genes recovered from
483 the metagenomic sequencing data. This *in silico* analysis indicated that primer sets comaA-D
484 (244f-659R), comaA (244F-659R), and comA-F/R matched recovered *amoA* genes, yet the issue
485 with unspecific product amplification and non-detects in some samples were persistent despite
486 additional efforts to optimize annealing temperatures, template concentration and PCR additives
487 (i.e., DMSO, BSA, magnesium, etc). Therefore, we developed new primer sets targeting the *amoB*
488 gene of comammox bacteria, specifically clade A comammox bacteria; the primary comammox
489 clade detected in our study and in wastewater treatment systems by recent studies (Chao et al.
490 2016, Pjevac et al. 2017, Roots et al. 2019, Spasov et al. 2019). The *amoB* gene exhibits
491 phylogenetic clustering consistent with that of *amoA* gene and places comammox bacteria in a
492 distinct cluster from all other ammonia oxidizers. Thus, we developed two primers sets targeting
493 the *amoB* gene of comammox bacteria with an expected product size of 114 and 337 bp. Both
494 primer sets resulted in high specificity (i.e., no unspecific amplification) (Figure 4C), high PCR
495 efficiency (~94%) (Figure 4D), and a clean melting curve (Figure 4E). Further, the qPCR bases
496 estimates of the ratio of comammox bacteria (determined using *amoB* specific primers) to total
497 ammonia oxidizers was highly correlated for the ratio of comammox bacteria to total ammonia
498 oxidizers determined using metagenomic data involving gene centric *de novo* assembly (Figure
499 4F) and mapping of metagenomic reads to reference genomes (Figure 4G). This suggests that the
500 newly designed primers are specific to clade A comammox bacteria and accurately capture their
501 abundance in complex nitrifying communities.

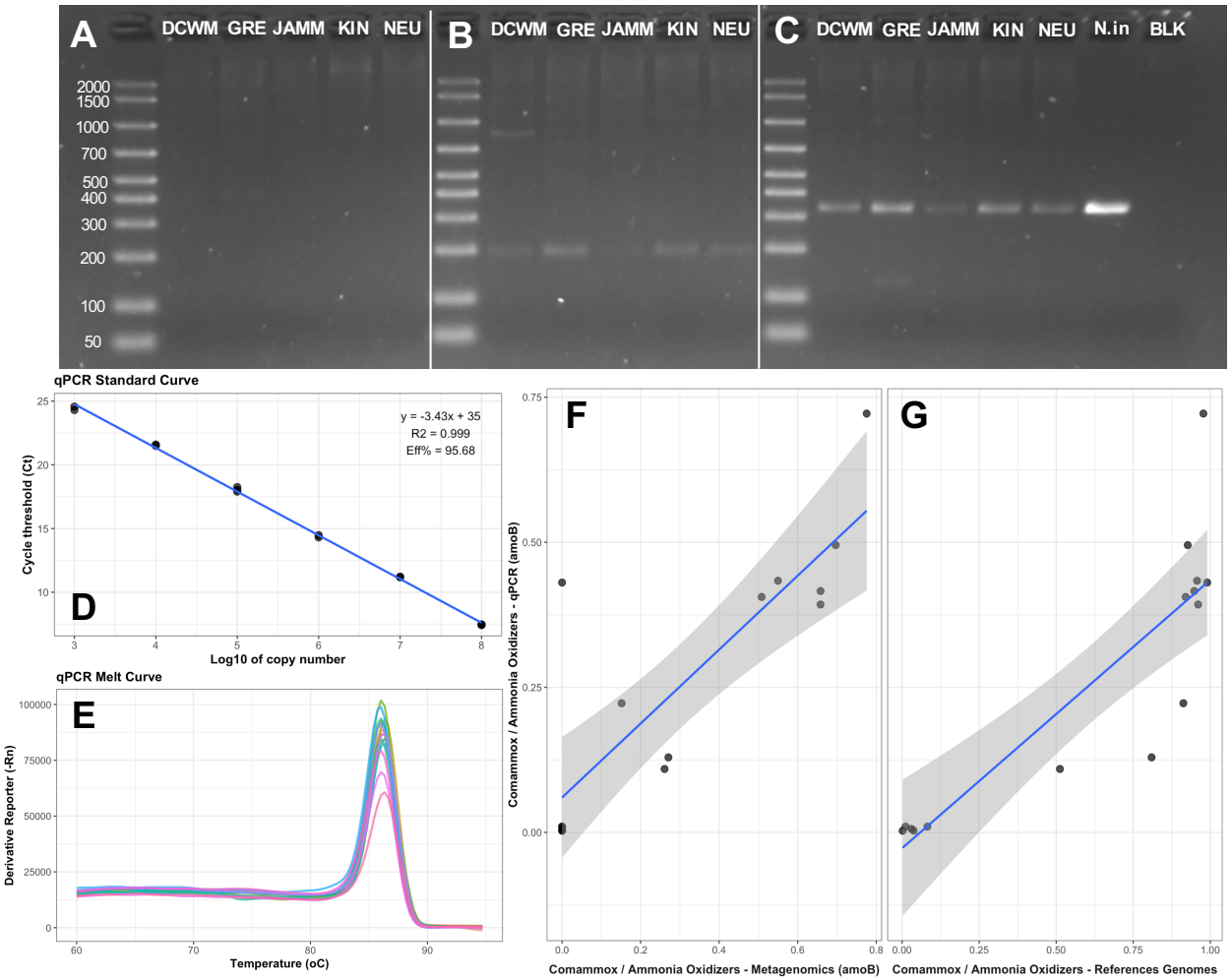


Figure 4: Previously published primer sets for the amoA gene of comammox bacteria, i.e. (A) comaA, and (B) Ntspa-amoA, result in non-detection and unspecific amplification samples from DCWM, GRE, JAMM, KIN, NEU where comammox bacteria were detected using metagenomics. In contrast the newly designed *amoB* specific primers (C) result in a single band formation of the right size for these samples which is consistent with that amplified from Ca Nitrospira inopinata (N.in). BLK stands for PCR negative control. The primers exhibit excellent PCR efficiency (D), demonstrate clean melting curves (E), and provide estimates of comammox bacterial abundance consistent with those seen in PCR-free metagenomics analyses (F, G).

3.5 Quantitative detection of nitrifying populations including comammox bacteria.

Initial assessment of qPCR results from the *amoA* assay for AOB underestimated AOB abundance when compared to the qPCR data targeting the 16S rRNA gene for AOB (Supplemental Figure S1). In contrast, 16S rRNA gene based qPCR estimates for AOB were consistent with those obtained using metagenomics (Supplemental Figure S1). This issue was also reported by Dechesne et al. (2016) (Dechesne et al. 2016) who indicated that a*moA* primer sets do not provide sufficient

510 coverage for AOB, while the 16S rRNA primer sets likely capture some non-AOB sequences as
511 well. Our analyses suggest that while *amoA* primer sets did not provide sufficient coverage for
512 AOB in our study, comparisons with metagenomic data indicated no evidence on unspecific non-
513 AOB detection with 16S rRNA gene primers. As a result, all subsequent measurements of AOB
514 were conducted using 16S rRNA gene based assays.
515
516 Canonical AOB and *Nitrospira* (including comammox and NOB) were detected in all systems
517 irrespective of the nitrogen removal process configuration. The relative abundance of both groups
518 ranges from 0.25 to 9% of total bacteria with an average of 1.71 and 1.65%, respectively, except
519 for *Nitrospira* in the attached growth of JAMM/JAMMSM which reached a maximum relative
520 abundance of 20% with an average of 8.56% (Figure 5A, B). For several systems, the abundance
521 of *Nitrospira* in proportion to AOB was significantly and consistently above what would be
522 expected if nitrification was being driven by AOB and NOB alone. These included the following
523 ND systems: DCWM, attached growth phase of JAMM (i.e., JAMMSM), KIN, SND system:
524 YAN, and PNA systems: attached growth phase of JAMS (i.e., JAMS). While comammox
525 bacteria were detected in most systems with higher proportional abundance of *Nitrospira* over
526 AOB (except for the PNA system: JAMS), the presence of comammox bacteria was not exclusive
527 to them (Figure 5B). For instance, comammox bacteria were also detected in GRE and NEU where
528 the abundance of *Nitrospira* and AOB were largely in line with expected proportions if nitrification
529 was primarily driven by AOB and NOB (Costa et al. 2006). Comammox bacteria were not detected
530 in four ND systems (ARM, BOA, NAN, YORM), one SND system (NAN), and the three PNA
531 systems (DUR/DURSM, JAMS/JAMSSM, YORS), which was consistent with metagenomic
532 observations. In all, comammox bacteria were detected in six ND systems (i.e., DCWM, GRE,
533 NEU, WIL, KIN, JAM/JAMSM) and one SND system (i.e., YAN). The abundance of comammox
534 bacteria in these systems ranged from 0.5% to ~3% of total bacteria in the suspended growth phase,
535 while it was as high (as high as ~17%) in the attached phase of the IFAS systems (i.e., JAMMSM).
536 In systems that included only suspended growth phase, comammox bacteria constituted between
537 20-70% of all ammonia oxidizers (i.e., AOB + comammox) for the duration of the study, with
538 abundances equal to that of AOB in two ND systems (i.e., GRE and NEU). The proportion of
539 comammox bacteria to that of total ammonia oxidizers varied significantly for two systems with

540 low dissolved oxygen, i.e., KIN (ND) and YAN (SND), from 1 to 70% and at times significantly
541 surpassed that of AOB**.**

542

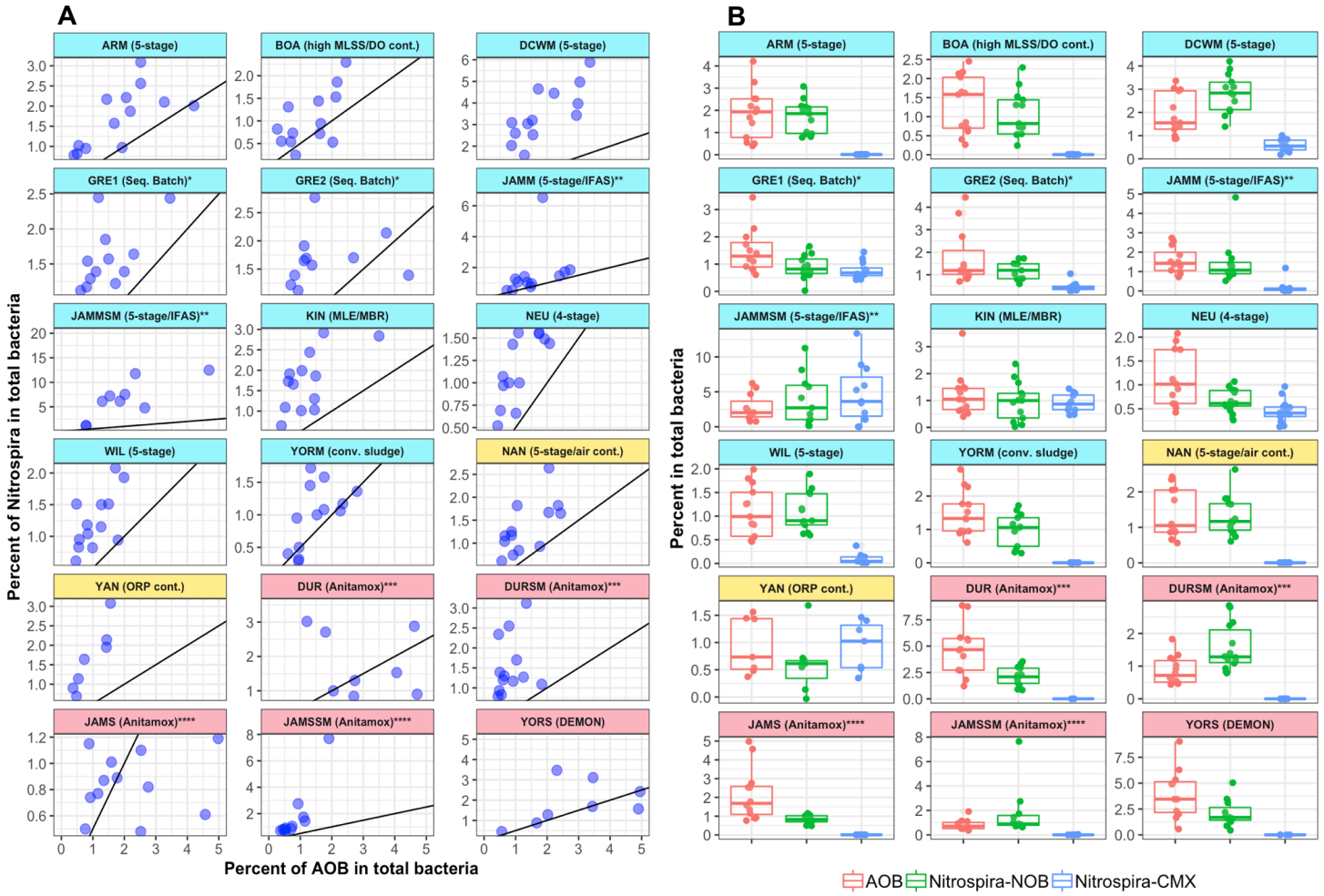


Figure 5: (A) Comparison of the abundance of Nitrospira (% of total bacteria) to that of AOB (% of total bacteria) indicates that several systems with higher proportional abundance of Nitrospira when compared to the theoretical estimates (black line) if nitrification was being driven by AOB and NOB. (B) The abundance of AOB, Nitrospira-NOB, and comammox bacteria as a proportional of total bacterial abundance indicates that comammox bacteria were (1) consistently detected in systems with metagenomic evidence of their presence, (2) their abundance was lower than AOB in suspended growth systems where they were detected and high in systems with attached growth phase. Asterisks levels indicate samples from the same system.

543 **3.6 Comammox bacterial abundance was strongly associated with solids retention time.**
544 Comammox bacteria were prevalent and temporally persistent in systems with long SRTs, despite
545 fluctuations in SRT levels within each system. The abundance of comammox bacteria as a
546 proportion of all ammonia oxidizers (i.e., comammox + AOB) increased with increasing median
547 SRT of the system (Figure 6A, 6B), with high abundances on the attached growth phase of the
548 IFAS systems at JAMM. Further, the mean ammonia removal was typically higher in systems with
549 comammox bacterial present compared to those without comammox bacteria (Wilcoxon rank sum
550 test, $p<0.0001$) (Figure 6C). And finally, while comammox bacterial abundance increased with
551 SRT (Figure 6F) (Spearman's $R = 0.52$, $p <0.0001$), the abundance of *Nitrospira*-NOB and AOB
552 did not exhibit significant association with SRT (Figure 6D, E). This suggests that increases in
553 comammox abundance may not be associated with a concomitant decrease in AOB or NOB
554 concentrations.
555
556 These data provide two potential insights into comammox bacterial relevance in nitrogen removal
557 systems and their potential competitive dynamics or lack thereof with canonical nitrifiers. First, as
558 stated earlier (section 3.3), comammox bacteria are preferentially enriched in systems with long
559 SRT's and systems with an attached phase component. While the prevalence in long SRT systems
560 may be explained by their slower net growth rates compared to canonical AOB and NOB (Kits et
561 al. 2017, Lawson and Lücker 2018), this does not provide basis for explaining their preferential
562 enrichment over AOB and NOB with increasing SRT. It is also important to note that the AOB
563 and NOB abundances were not associated with a concomitant increase in comammox bacterial
564 abundance suggests that these bacteria may occupy exclusive niches within the nitrifying
565 consortium. For instance, AOB and comammox may occupy independent niches at different
566 ammonia concentrations due to different affinity levels for ammonia (Kits et al. 2017, Lawson and
567 Lücker 2018) which may allow for their co-existence. In fact, in three ND systems (i.e., DCWM,
568 NEU and YAN) and the IFAS ND system (i.e., JAMMSM) the concentrations of AOB and
569 comammox were positively correlated ($R = 0.53$ (DCWM), 0.41 (NEU), 0.91 (YAN), 0.60
570 (JAMMSM), $p<0.01$), suggesting the potential for some level of cooperation between the two.
571 Further, like other lineage II *Nitrospira*, comammox bacteria demonstrate potential to exhibit
572 diverse metabolic capacities (Daims et al. 2016, Lawson and Lücker 2018) including genes
573 involved in urea uptake and urease for conversion of urea to ammonia (Daims et al. 2016, Koch et

al. 2015, Lawson and Lücker 2018). Thus, it is plausible that the combination of greater metabolic diversity (e.g., ability to utilize urea) and higher affinities for ammonia of comammox bacteria compared to that of AOB, may ensure that comammox bacteria may have preferential access to ammonia (or other electron donors) made available via increased biomass decay products at longer SRTs. This could likely explain the increase in the abundance of comammox bacteria and not AOB or NOB with increasing SRT.

The second insight is that systems with comammox bacteria may have higher ammonia removal efficiencies compared to nitrogen removal systems where they are absent. It is important to note that this statistically significant observation is confounded by the fact that systems with comammox bacteria also exhibited higher SRT. Thus, it is difficult to disentangle the role of comammox bacteria from that of SRT alone in explaining the differences in ammonia removal efficiency. Finally, we did not find any other significant correlations between comammox bacterial concentrations and proportions (of total bacteria and ammonia oxidizers) with any of the other measured process parameters except including loading rates, pH, temperature, and dissolved oxygen (DO) concentrations. The lack of association with DO concentrations is in contrast with previous studies (Beach and Noguera 2019, Roots et al. 2019). In fact, comammox bacteria were present and their abundances stable in systems with DO less than 1 mg/l (i.e., YAN, NEU) as well as those with DO levels significantly in excess of 2 mg/l (i.e., GRE, JAMM, WIL, KIN). It is important to note that correlations of comammox bacterial abundance and proportions with process parameters were estimated on a system-by-system basis, with a maximum of 12 data points per system. Thus, it is feasible that our correlation efforts were limited by the richness of our dataset on an individual system basis, where multiple factors may together play a role in influencing microbial community composition. Thus, while this does not eliminate the possibility that low DO levels (or other process parameters) impact comammox bacteria, it suggests that SRT plays a more prominent role compared to other process parameters in the selection of comammox bacteria in full-scale mainstream nitrogen removal systems.

601

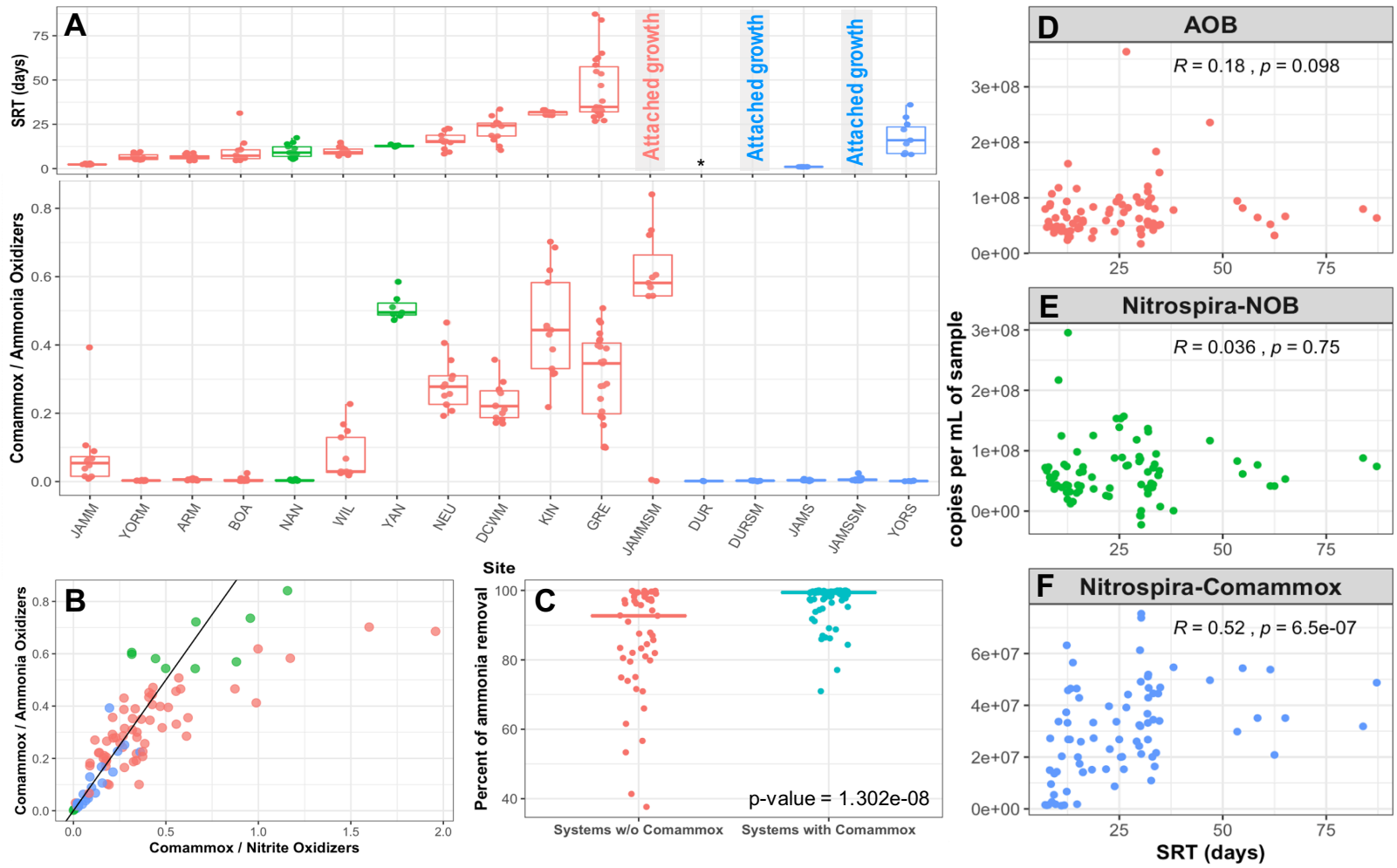


Figure 6: Increase in solids retention time (A) was associated with comammox bacteria constituting a greater proportion of all ammonia oxidizers (B). (C) Systems with comammox bacteria exhibited greater ammonia removal efficiencies compared to systems without comammox bacteria. This comparison does not include PNA systems. In contrast to AOB (D) and NOB (E), whose abundance was not associated with SRT, comammox bacterial abundance (F) was significantly and positively correlated with system SRT.

4.0 Conclusions

- Comammox bacteria were not prevalent in side-stream PNA systems, including single stage systems that contain attached growth and suspended phase components. This is likely to be associated with the high ambient ammonia concentrations in PNA systems.
- All comammox bacteria detected in full-scale systems belonged to clade A comammox bacteria and are closely associated with *Ca* Nitrospira nitrosa. This study provides a novel primer set and qPCR assay targeting the *amoB* gene of clade A comammox bacteria.
- Comammox bacteria were prevalent in full-scale mainstream nitrogen removal systems with long SRT's and/or systems with attached growth components. This finding is consistent with estimates of slower growth rates for comammox bacteria compared to canonical AOB and NOB.
- Increases in comammox bacterial abundance in systems with sufficient SRT and/or attached growth phase were not associated with a concomitant decrease in the abundance of canonical AOB or NOB indicating that they may occupy niche independent from that of canonical nitrifiers within complex nitrifying communities.
- We found no significant associations between DO concentrations and comammox presence/absence or concentration in this study. While this does not eliminate the possibility that low DO levels favor comammox bacteria in wastewater treatment systems, it suggests that SRT is the key variable driving the prevalence of comammox bacteria.

5.0 Acknowledgements

This research was supported by the Water Environment Reuse Foundation (Grant UR416) and The National Science Foundation (Award number: 1703089). The authors also acknowledge the operational personnel for assistance with sampling and process data sharing. The authors would also like to thank Prof. Kartik Chandran (Columbia University) and Prof. Holger Daims (University of Vienna) and Dr. Petra Pjevac (University of Vienna) for providing DNA extracts of *Nitrosomonas europaea* and *Ca* Nitrospira inopinata, respectively, for preparation for qPCR standards.

633 **6.0 References**

634 Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment
635 search tool. Journal of Molecular Biology 215(3), 403-410.
636 Annavajhala, M.K., Kapoor, V., Santo-Domingo, J. and Chandran, K. (2018) Comammox
637 functionality identified in diverse engineered biological wastewater treatment systems.
638 Environmental Science & Technology Letters 5(2), 110-116.
639 Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S. and Ogata, H.
640 (2019) KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive
641 score threshold. bioRxiv, 602110.
642 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M.,
643 Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler,
644 G., Alekseyev, M.A. and Pevzner, P.A. (2012) SPAdes: A new genome assembly algorithm
645 and its applications to single-cell sequencing. Journal of Computational Biology 19(5), 455-
646 477.
647 Bartelme, R.P., McLellan, S.L. and Newton, R.J. (2017) Freshwater recirculating aquaculture
648 system operations drive biofilter bacterial community shifts around a stable nitrifying
649 consortium of ammonia-oxidizing archaea and comammox nitrospira. Frontiers in
650 Microbiology 8(101).
651 Beach, N.K. and Noguera, D.R. (2019) Design and assessment of species-level qPCR primers
652 targeting comammox. Frontiers in Microbiology 10(36).
653 Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina
654 sequence data. Bioinformatics 30(15), 2114-2120.
655 Buchfink, B., Xie, C. and Huson, D.H. (2015) Fast and sensitive protein alignment using
656 DIAMOND. Nature Methods 12(1), 59-60.
657 Camejo, P.Y., Santo Domingo, J., McMahon, K.D. and Noguera, D.R. (2017) Genome-enabled
658 insights into the ecophysiology of the comammox bacterium "*Candidatus* Nitrospira
659 nitrosa". mSystems 2(5), e00059-00017.
660 Campbell, B.J., Yu, L., Heidelberg, J.F. and Kirchman, D.L. (2011) Activity of abundant and
661 rare bacteria in a coastal ocean. Proceedings of the National Academy of Sciences 108(31),
662 12776.
663 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J.,
664 Fierer, N. and Knight, R. (2011) Global patterns of 16S rRNA diversity at a depth of millions
665 of sequences per sample. Proceedings of the National Academy of Sciences 108(Supplement
666 1), 4516-4522.
667 Chao, Y., Mao, Y., Yu, K. and Zhang, T. (2016) Novel nitrifiers and comammox in a full-scale
668 hybrid biofilm and activated sludge reactor revealed by metagenomic approach. Applied
669 Microbiology and Biotechnology 100(18), 8225-8237.
670 Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro,
671 A., Kuske, C.R. and Tiedje, J.M. (2014) Ribosomal Database Project: data and tools for high
672 throughput rRNA analysis. Nucleic Acids Research 42(Database issue), D633-D642.
673 Costa, E., Pérez, J. and Kreft, J.U. (2006) Why is metabolic labour divided in nitrification?
674 Trends in Microbiology 14(5), 213-219.
675 Daims, H., Lebedeva, E.V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N.,
676 Palatinszky, M., Vierheilig, J., Bulaev, A., Kirkegaard, R.H., von Bergen, M., Rattei, T.,
677 Bendinger, B., Nielsen, P.H. and Wagner, M. (2015) Complete nitrification by Nitrospira
678 bacteria. Nature 528(7583), 504-509.

679 Daims, H., Lücker, S. and Wagner, M. (2016) A new perspective on microbes formerly known
680 as Nitrite-Oxidizing bacteria. Trends in Microbiology 24(9), 699-712.
681 Dechesne, A., Musovic, S., Palomo, A., Diwan, V. and Smets, B.F. (2016) Underestimation of
682 ammonia-oxidizing bacteria abundance by amplification bias in amoA-targeted qPCR.
683 Microbial Biotechnology 9(4), 519-524.
684 Eddy, S.R. (2011) Accelerated profile HMM searches. PLOS Computational Biology 7(10),
685 e1002195.
686 Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high
687 throughput. Nucleic Acids Research 32(5), 1792-1797.
688 Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L. and Delmont,
689 T.O. (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ
690 3, e1319.
691 Fan, X.Y., Gao, J.F., Pan, K.L., Li, D.C. and Dai, H.H. (2017) Temporal dynamics of bacterial
692 communities and predicted nitrogen metabolism genes in a full-scale wastewater treatment
693 plant. RSC Advances 7(89), 56317-56327.
694 Fowler, S.J., Palomo, A., Dechesne, A., Mines, P.D. and Smets, B.F. (2018) Comammox
695 Nitrospira are abundant ammonia oxidizers in diverse groundwater-fed rapid sand filter
696 communities. Environmental Microbiology 20(3), 1002-1015.
697 Gonzalez-Martinez, A., Rodriguez-Sanchez, A., van Loosdrecht, M.C.M., Gonzalez-Lopez, J.
698 and Vahala, R. (2016) Detection of comammox bacteria in full-scale wastewater treatment
699 bioreactors using tag-454-pyrosequencing. Environmental Science and Pollution Research
700 23(24), 25501-25511.
701 Graham, D., Knapp, C., Van Vleck, E., Bloor, K., Lane, T. and Graham, C. (2007) Experimental
702 demonstration of chaotic instability in biological nitrification. The ISME Journal 1(5), 385-
703 393.
704 Hermansson, A. and Lindgren, P.E. (2001) Quantification of ammonia-oxidizing bacteria in
705 arable soil by Real-Time PCR. Applied and Environmental Microbiology 67(2), 972-976.
706 Hu, H.W. and He, J.Z. (2017) Comammox - a newly discovered nitrification process in the
707 terrestrial nitrogen cycle. Journal of Soils and Sediments 17(12), 2709-2717.
708 Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W. and Hauser, L.J. (2010)
709 Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC
710 Bioinformatics 11, 119.
711 Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T. and Aluru, S. (2018) High
712 throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries.
713 Nature Communications 9(1), 5114.
714 Juretschko, S., Timmermann, G., Schmid, M., Schleifer, K.H., Pommerening-Roser, A., Koops,
715 H.P. and Wagner, M. (1998) Combined molecular and conventional analyses of nitrifying
716 bacterium diversity in activated sludge: Nitrosococcus mobilis and Nitrospira-like bacteria as
717 dominant populations. Applied and Environmental Microbiology 64(8), 3042-3051.
718 Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M. (2016) KEGG as a
719 reference resource for gene and protein annotation. Nucleic Acids Research 44(D1), D457-
720 D462.
721 Kang, D.D., Froula, J., Egan, R. and Wang, Z. (2015) MetaBAT, an efficient tool for accurately
722 reconstructing single genomes from complex microbial communities. PeerJ 3, e1165.

723 Kits, K.D., Sedlacek, C.J., Lebedeva, E.V., Han, P., Bulaev, A., Pjevac, P., Daebeler, A.,
724 Romano, S., Albertsen, M., Stein, L.Y., Daims, H. and Wagner, M. (2017) Kinetic analysis
725 of a complete nitrifier reveals an oligotrophic lifestyle. Nature 549, 269-272.
726 Klotz, M.G. and Stein, L.Y. (2008) Nitrifier genomics and evolution of the nitrogen cycle.
727 FEMS Microbiol Letters 278(2), 146-156.
728 Koch, H., Lücker, S., Albertsen, M., Kitzinger, K., Herbold, C., Spieck, E., Nielsen, P.H.,
729 Wagner, M. and Daims, H. (2015) Expanded metabolic versatility of ubiquitous nitrite-
730 oxidizing bacteria from the genus Nitrospira. Proceedings of the National Academy of
731 Sciences 112(36), 11371-11376.
732 Kowalchuk, G.A. and Stephen, J.R. (2001) Ammonia-oxidizing bacteria: a model for molecular
733 microbial ecology. Annual Review of Microbiology 55, 485-529.
734 Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. Nature
735 Methods 9, 357-359.
736 Lawson, C.E. and Lücker, S. (2018) Complete ammonia oxidation: an important control on
737 nitrification in engineered ecosystems? Current Opinion in Biotechnology 50, 158-165.
738 Li, H. and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows–Wheeler
739 transform. Bioinformatics 26(5), 589-595.
740 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G.,
741 Durbin, R. and Subgroup, G.P.D.P. (2009) The sequence alignment/map format and
742 SAMtools. Bioinformatics 25(16), 2078-2079.
743 Matsen, F.A., Kodner, R.B. and Armbrust, E.V. (2010) pplacer: linear time maximum-likelihood
744 and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC
745 Bioinformatics 11, 538.
746 McIlroy, S.J., Saunders, A.M., Albertsen, M., Nierychlo, M., McIlroy, B., Hansen, A.A., Karst,
747 S.M., Nielsen, J.L. and Nielsen, P.H. (2015) MiDAS: the field guide to the microbes of
748 activated sludge. Database 2015.
749 Miller, C., Baker, B., Thomas, B., Singer, S. and Banfield, J. (2011) EMIRGE: reconstruction of
750 full-length ribosomal genes from microbial community short read sequencing data. Genome
751 Biology 12(5), R44.
752 Nurk, S., Meleshko, D., Korobeynikov, A. and Pevzner, P.A. (2017) metaSPAdes: a new
753 versatile metagenomic assembler. Genome Research 27(5), 824-834.
754 Orellana, L.H., Hatt, J.K., Iyer, R., Chourey, K., Hettich, R.L., Spain, J.C., Yang, W.H., Chee-
755 Sanford, J.C., Sanford, R.A., Löffler, F.E. and Konstantinidis, K.T. (2018) Comparing DNA,
756 RNA and protein levels for measuring microbial activity in nitrogen-amended soils. bioRxiv,
757 466680.
758 Palomo, A., Jane Fowler, S., Gulay, A., Rasmussen, S., Sicheritz-Ponten, T. and Smets, B.F.
759 (2016) Metagenomic analysis of rapid gravity sand filter microbial communities suggests
760 novel physiology of Nitrospira spp. The ISME Journal 10(11), 2569-2581.
761 Palomo, A., Pedersen, A.G., Fowler, S.J., Dechesne, A., Sicheritz-Pontén, T. and Smets, B.F.
762 (2018) Comparative genomics sheds light on niche differentiation and the evolutionary
763 history of comammox Nitrospira. The ISME Journal 12(7), 1779-1793.
764 Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A. and
765 Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny
766 substantially revises the tree of life. Nature Biotechnology 36, 996-1004.

767 Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P. and Tyson, G.W. (2015) CheckM:
768 assessing the quality of microbial genomes recovered from isolates, single cells, and
769 metagenomes. Genome Research 25(7),1043-55.
770 Pericard, P., Dufresne, Y., Couderc, L., Blanquart, S. and Touzet, H. (2018) MATAM:
771 reconstruction of phylogenetic marker genes from short sequencing reads in metagenomes.
772 Bioinformatics 34(4):585-591.
773 Pinto, A.J., Marcus, D.N., Ijaz, U.Z., Bautista-de lose Santos, Q.M., Dick, G.J. and Raskin, L.
774 (2015) Metagenomic Evidence for the Presence of Comammox Nitrospira-Like Bacteria in a
775 Drinking Water System. mSphere 1(1), e00054-00015.
776 Pjevac, P., Schauberger, C., Poghosyan, L., Herbold, C.W., van Kessel, M.A.H.J., Daebeler, A.,
777 Steinberger, M., Jetten, M.S.M., Lücker, S., Wagner, M. and Daims, H. (2017) AmoA-
778 targeted polymerase chain reaction primers for the specific detection and quantification of
779 comammox nitrospira in the environment. Frontiers in Microbiology 8(1508).
780 Poghosyan, L., Koch, H., Lavy, A., Frank, J., van Kessel, M.A.H.J., Jetten, M.S.M., Banfield,
781 J.F. and Lücker, S. (2019) Metagenomic recovery of two distinct comammox Nitrospira from
782 the terrestrial subsurface. Environmental Microbiology.
783 Price, M.N., Dehal, P.S. and Arkin, A.P. (2010) Fasttree 2 – approximately maximum-likelihood
784 trees for large alignments. PLoS ONE 5(3): e9490.
785 Prosser, J.I. and Nicol, G.W. (2008) Relative contributions of archaea and bacteria to aerobic
786 ammonia oxidation in the environment. Environmental Microbiology 10(11), 2931-2941.
787 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner,
788 F.O. (2013) The SILVA ribosomal RNA gene database project: Improved data processing
789 and web-based tools. Nucleic Acids Research 41, 590-596.
790 RCoreTeam (2014) R: A language and environment for statistical computing R foundation for
791 statistical computing, Vienna, Austria.
792 Roots, P., Wang, Y., Rosenthal, A.F., Griffin, J.S., Sabba, F., Petrovich, M., Yang, F., Kozak,
793 J.A., Zhang, H. and Wells, G.F. (2019) Comammox Nitrospira are the dominant ammonia
794 oxidizers in a mainstream low dissolved oxygen nitrification reactor. Water Research 157,
795 396-405.
796 Rotthauwe, J.H., Witzel, K.P. and Liesack, W. (1997) The ammonia monooxygenase structural
797 gene amoA as a functional marker: molecular fine-scale analysis of natural ammonia-
798 oxidizing populations. Applied and Environmental Microbiology 63(12), 4704-4712.
799 Saunders, A.M., Albertsen, M., Vollertsen, J. and Nielsen, P.H. (2016) The activated sludge
800 ecosystem contains a core community of abundant organisms. The ISME Journal 10(1), 11-
801 20.
802 Siripong, S. and Rittmann, B.E. (2007) Diversity study of nitrifying bacteria in full-scale
803 municipal wastewater treatment plants. Water Research 41(5), 1110-1120.
804 Spasov, E., Tsuji, J.M., Hug, L.A., Doxey, A.C., Sauder, L.A., Parker, W.J. and Neufeld, J.D.
805 (2019) Comammox bacteria are dominant ammonia oxidizers in rotating biological
806 contactors of a municipal wastewater treatment plant. bioRxiv, 529826.
807 Stahl, D.A., Torre, J.R.d.l., SM, A. and JI, P. (2012) Physiology and diversity of ammonia-
808 oxidizing archaea urease activity in neutrophilic autotrophic ammonia-oxidizing bacteria
809 isolated from acid soils. Annual Review of Microbiology 66(1), 83-101.
810 Stamatakis, A. (2014) RAxML Version 8: A tool for phylogenetic analysis and post-analysis of
811 large phylogenies. Bioinformatics 30(9):1312-1313.

812 Tatari, K., Musovic, S., Gülay, A., Dechesne, A., Albrechtsen, H.-J. and Smets, B.F. (2017)
813 Density and distribution of nitrifying guilds in rapid sand filters for drinking water
814 production: Dominance of Nitrospira spp. Water Research 127, 239-248.
815 van Kessel, M.A.H.J., Speth, D.R., Albertsen, M., Nielsen, P.H., Op den Camp, H.J.M., Kartal,
816 B., Jetten, M.S.M. and Lücker, S. (2015) Complete nitrification by a single microorganism.
817 Nature 528(7583), 555-559.
818 Wang, M., Huang, G., Zhao, Z., Dang, C., Liu, W. and Zheng, M. (2018) Newly designed primer
819 pair revealed dominant and diverse comammox amoA gene in full-scale wastewater
820 treatment plants. Bioresource Technology 270, 580-587.
821 Wang, Q., Garrity, G., Tiedje, J. and Cole, J. (2007) Naive Bayesian classifier for rapid
822 assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental
823 Microbiology 73(16), 5261-5267.
824 Wang, Y., Ma, L., Mao, Y., Jiang, X., Xia, Y., Yu, K., Li, B. and Zhang, T. (2017) Comammox
825 in drinking water systems. Water Research 116, 332-341.
826 Wick, R.R., Judd, L.M., Gorrie, C.L. and Holt, K.E. (2017) Unicycler: Resolving bacterial
827 genome assemblies from short and long sequencing reads. PLOS Computational Biology
828 13(6), e1005595.
829 Wickham, H. (2009). ggplot2: elegant graphics for data analysis. New York: Springer.
830 Wu, L., Ning, D., Zhang, B., Li, Y., Zhang, P., Shan, X., Zhang, Q., Brown, M., Li, Z., Van
831 Nostrand, J.D., Ling, F., Xiao, N., Zhang, Y., Vierheilig, J., Wells, G.F., Yang, Y., Deng, Y.,
832 Tu, Q., Wang, A., Global Water Microbiome Consortium, Zhang, T., He, Z., Keller, J.,
833 Nielsen, P.H., Alvarez, P.J.J., Criddle, C.S., Wagner, M., Tiedje, J.M., He, Q., Curtis, T.P.,
834 Stahl, D.A., Alvarez-Cohen, L., Rittmann, B.E., Wen, X. and Zhou, J. (2019) Global
835 diversity and biogeography of bacterial communities in wastewater treatment plants. Nature
836 Microbiology 4, 1183-1195.
837 Xia, F., Wang, J.G., Zhu, T., Zou, B., Rhee, S.K. and Quan, Z.X. (2018) Ubiquity and diversity
838 of complete ammonia oxidizers (comammox). Applied and Environmental Microbiology
839 84(24), e01390-01318.