Comammox bacteria are dominant ammonia oxidizers in tertiary rotating biological contactors of a municipal wastewater treatment plant

Emilie Spasov[1], Jackson M. Tsuji[1], Laura A. Hug[1], Andrew C. Doxey[1], Laura A. Sauder[1], Wayne J. Parker[2], Josh D. Neufeld[1*]

[1]Department of Biology, University of Waterloo, Waterloo, Ontario, Canada

[2]Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, Ontario, Canada

*Corresponding author: Department of Biology, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L 3G1, Canada. Tel. +1 519-888-4567; Fax +1 519-746-0614. E-mail: jneufeld@uwaterloo.ca

Running title: Comammox bacteria dominate wastewater biofilm

Keywords: complete ammonia oxidation, ammonia-oxidizing archaea, wastewater, nitrification, biofilm, metagenomics, *amoA*

Abstract

Nitrification, the oxidation of ammonia to nitrate via nitrite, is an important process in wastewater treatment plants (WWTPs), yet little is known about the abundance and diversity of the recently discovered complete ammonia-oxidizing (comammox) bacteria in these systems. This study investigated ammonia oxidizing bacteria (AOB), including comammox bacteria, and ammonia oxidizing archaea (AOA) in biofilm samples collected from full scale rotating biological contactors (RBCs) that are employed in the tertiary treatment system of a municipal WWTP in Guelph, Ontario, Canada. Temporal and spatial diversity and abundance of RBC microbial communities were analyzed using quantitative PCR (qPCR) and metagenomics. Metagenomic taxonomic profiling and qPCR revealed that *Nitrospira* species dominated most RBC biofilm samples, and functional profiling with ammonia monooxygenase (*amoA*) and nitrite oxidoreductase (*nxrB*) gene markers indicated that comammox-associated *Nitrospira* spp. represented the most abundant group of ammonia-oxidizing microorganisms. Genome bins recovered from metagenomic data indicated that multiple populations of comammox bacteria were present including novel comammox taxa with distinct spatial and temporal distributions. This study indicates that RBCs are valuable systems for investigating environmental factors governing distributions and activity of comammox bacteria, other AOB, AOA, and NOB. With these baseline data, future research will leverage the unique design and arrangement of RBCs for investigating the distinct niches occupied by these bacterial and archaeal nitrifiers.

Introduction

Municipal wastewater contains ammonia that is removed by wastewater treatment plants (WWTPs) in order to prevent eutrophication, oxygen depletion, and toxicity to aquatic animals in receiving waters. Nitrification involves the oxidation of ammonia to nitrate via nitrite in two enzymatic steps that were historically thought to be mediated by distinct groups of microorganisms (Stein and Klotz, 2016). Aerobic ammonia oxidation was thought to be conducted solely by ammonia-oxidizing bacteria (AOB) and ammonia-oxidizing archaea (AOA), and nitrite oxidation was associated with strict nitrite-oxidizing bacteria (NOB) (Daims *et al.*, 2016; Könneke *et al.*, 2005; Purkhold *et al.*, 2000). The existence of microorganisms capable of catalyzing both steps of nitrification in the process of complete ammonia oxidation (comammox) was suggested over a decade ago, with a prediction that such bacteria would be relatively slow growing and inhabit biofilms that were exposed to relatively low ammonia concentrations (Costa *et al.*, 2006). Recently, these predictions were confirmed by the discovery of *Nitrospira* species capable of catalyzing comammox (Daims *et al.*, 2015; van Kessel *et al.*, 2015). All comammox bacteria reported to date belong to lineage II of the genus *Nitrospira*. Two clades of comammox bacteria have been described (clades A and B), based on ammonia monooxygenase (*amoA*) gene phylogeny, with existing cultivated species of comammox bacteria belonging to clade A (van Kessel *et al.*, 2015; Daims *et al.*, 2015).

Little is known about the abundance and diversity of comammox bacteria in engineered aquatic environments. First detected in water treatment system metagenomes (Daims *et al.*, 2015; van Kessel *et al.*, 2015), most comammox bacteria detected in WWTPs to date belong to comammox clade A (Pjevac *et al.*, 2017; Daims *et al.*, 2015; Xia *et al.*, 2018; Zhao *et al.*, 2018) and typically do not exceed abundances reported for AOA and other AOB (Pjevac *et al.*, 2017;

Chao *et al.*, 2016; Fan *et al.*, 2017; Pan *et al.*, 2018). However, comammox *amoA* gene abundances have recently been reported to outnumber those of AOB in activated sludge samples from several WWTPs (Wang *et al.*, 2018; Zhao *et al.*, 2018). Additionally, the high abundance of comammox *amoA* transcripts detected in activated sludge suggests that comammox bacteria have an active role in nitrification in these systems (Zheng *et al.*, 2019). Comammox bacteria have also been found in engineered environments with low ammonia concentrations, such as drinking water treatment plants (Pjevac *et al.*, 2017; Pinto *et al.*, 2015; Palomo *et al.*, 2016; Wang *et al.*, 2017). Previous studies examining comammox bacterial distributions in WWTPs have focused on activated sludge secondary treatment systems (Gonzalez-Martinez *et al.*, 2016; Fan *et al.*, 2017; Chao *et al.*, 2016; Pan *et al.*, 2018; Wang *et al.*, 2018; Zhao *et al.*, 2018). To date, no studies have investigated comammox bacteria in tertiary wastewater treatment systems.

Rotating biological contactors (RBCs; Figure 1A) represent a tertiary treatment system of the Guelph WWTP (Ontario, Canada). Comprising a total nitrifying biofilm surface area of 440,000 m^2 when fully operational (Guelph Wastewater Treatment Master Plan, 2009), the RBCs are organized into four “trains” that are each composed of eight individual RBC “stages” (Figure 1B). Continual rotation in and out of post-aeration basin wastewater allows for alternating exposure of RBC biofilm to wastewater and air, resulting in decreasing ammonia concentration along the RBC flowpath due to aerobic nitrification, as described in Sauder *et al.* (2012). Because of the relatively low ammonia concentrations entering the RBCs (i.e. compared to aeration basins), and their fixed-film design, these RBCs present a valuable opportunity to study the distributions and diversity of nitrifiers in relation to ammonia concentration gradients. The AOA and AOB communities in these RBCs were previously characterized, demonstrating that the abundance of an AOA population increased along the RBC flowpath (Sauder *et al.*,

2012). AOA from this population were later cultivated and named *Candidatus* Nitrosocosmicus exaquare (Sauder *et al.*, 2017). However, inconsistencies in activity assessments with differential inhibitors and the abundance of *Nitrospira* cells that actively assimilated bicarbonate autotrophically indicated the potential for comammox bacteria (Sauder *et al.*, 2017). The discovery of comammox bacteria has necessitated a re-evaluation of the ecology of nitrification in wastewater systems, including the Guelph RBCs.

Based on inconsistencies in previous data, and due to the predicted low ammonia niche and biofilm growth of comammox bacteria (Costa *et al.*, 2006; Daims *et al.*, 2015), we hypothesized that comammox bacteria are present in the Guelph WWTP RBCs, adding additional complexity to the classical interpretation of RBC microbiology with respect to nitrification. We also predicted that the relative abundance of comammox bacteria would increase as ammonia concentrations decreased along the RBC flowpath. To test these hypotheses, we assessed the relative abundance and diversity of comammox bacteria, in relation to AOA and traditional AOB, throughout the tertiary treatment system using a combination of qPCR and metagenomic sequencing of RBC samples. We also analyzed the temporal stability of RBC biofilm communities in three time points in 2010 and one time point in 2016. Furthermore, through assembly and binning of metagenome sequences, we confirmed whether genome bins representing potential RBC comammox bacterial populations contained the necessary genetic complements to mediate complete ammonia oxidation, and we assessed the taxonomic diversity and novelty of these populations compared to known comammox *Nitrospira*.

Methods

Sampling

Using ethanol-cleaned spatulas, biofilm was sampled from the RBC trains in Guelph, Ontario, Canada (Figure 1A, Figure 1B). Samples were stored on dry ice until delivered to the lab and were kept at -70°C until analyzed. Biofilm was sampled in February, June, and September in 2010 and in October 2016. In June and September 2010, all eight stages of both the northeast (NE) and southeast (SE) trains were sampled, and eight stages of the NE train were sampled in February 2010 (Sauder *et al.*, 2012); only the three temporal samples from the NE RBC 1 and RBC 8 were used in this current study (Table S1). All RBCs from all four trains were sampled in 2016, except for RBC 1 of the SE train and RBC 2 of the southwest (SW) train, which were not operational. Water samples were also collected from RBC 1 and RBC 8 and kept on ice and then frozen at -20°C until chemical analysis.

Water chemistry

Total ammonia ($NH_3 + NH_4^+$) was measured fluorometrically using orthophthaldialdehyde (OPA) reagent (Holmes *et al.*, 1999) according to the method by Poulin and Pelletier (2007), with minor modifications. Volumes of 100 µL of sample and 200 µL of OPA working reagent were added to a 96-well opaque flat-bottomed plate. Plates were incubated for four hours in the dark before being measured. Nitrite and nitrate concentrations were assessed colorimetrically using the Greiss reagent as described by Miranda *et al.* (2001). All samples were measured as technical duplicates at 360 nm excitation, 465 nm emission (ammonia), and 550 nm (nitrite/nitrate) using a FilterMax F5 Multi-Mode Microplate reader (Molecular Devices, San Jose, CA, USA). Measurements reflect total ammonia as nitrogen, and nitrite and nitrate as

nitrogen. Monthly and yearly ammonia concentrations (measured as total ammonia as nitrogen) of secondary effluent (RBC influent) and water from the southeast RBC train were compiled for the years 2010 to 2017. Data were obtained from WWTP operators. Plant influent and effluent data were retrieved from the Guelph WWTP annual reports for 2010 to 2017, available at https://guelph.ca/living/environment/water/wastewater/.

DNA extractions

All extractions were done with the PowerSoil DNA Isolation Kit (Mo Bio, Carlsbad, CA, USA) according to the manufacturer's instructions, except that bead beating was conducted with a FastPrep-24 (MP Biomedicals, Santa Ana, CA, USA) at 5.5 m/s for 45 s. Genomic DNA was visualized on a 1% agarose gel and quantified using a NanoDrop 2000 (Thermo Scientific, Waltham, MA, USA) and Qubit dsDNA high sensitivity assay kit (Thermo Scientific).

Extracted DNA from RBC 1 and 8 biofilm samples of the northeast train from 2010 was used for quantitative PCR (qPCR) and metagenomic sequencing (Table S1). The 2010 biofilm samples used were the same samples collected and analyzed by qPCR previously (Sauder *et al.*, 2012), but DNA extractions were repeated while processing the 2016 samples. The DNA extracts from all RBCs within the four trains sampled in 2016 were analyzed with qPCR, whereas DNA from RBCs 1 and 8 alone were analyzed with metagenomic sequencing (Table S1).

Quantitative PCR

Archaeal and bacterial 16S rRNA genes were quantified using the primers 771F/957R (Ochsenreiter *et al.*, 2003) and 341F/518R (Muyzer *et al.*, 1993), respectively (Table S2). Quantification of AOB *amoA* genes was carried out using the primers amoA1F/amoA2R

(Rotthauwe *et al.*, 1997) (Table S2). Comammox clade A and clade B were amplified using equimolar primer mixes of comaA-244f (a-f) and comaA-659r (a-f), and comaB-244f (a-f) and comaB-659r (a-f), respectively (Pjevac *et al.*, 2017) (Table S2). Primers for each of comammox clades A and B were initially tested in end-point PCR, but because clade B primers produced no amplicons (data not shown), subsequent qPCR was performed with clade A primers only. All qPCR amplifications were carried out as technical duplicates on a CFX96 Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA). Additional details on thermal cycling conditions and standards are included in the supplemental methods.

Metagenomic sequencing and analysis

DNA shearing, library preparation, and sequencing was performed at The Centre for Applied Genomics (TCAG) in Toronto, Ontario. Extracted DNA was quantified by Qubit dsDNA High Sensitivity assay (Thermo Scientific, Waltham, MA, USA). Subsequently, 500 ng of input DNA was sheared into ~550 bp fragments using a Covaris LE220 Focused-ultrasonicator (Covaris, Woburn, MA, USA). Library preparation was done with the TruSeq PCR-free Library Prep Kit (Illumina, San Diego, CA, USA). Paired-end sequencing (2x250 bases) was performed on a HiSeq 2500 (Illumina) using the HiSeq Rapid SBS Kit v2 (500 cycle) (Illumina), resulting in a total of ~250 million paired-end reads with an average of 18.1 million paired-end reads per sample (Table S3).

Quality trimming and removal of adaptor sequences was performed using AdapterRemoval version 2.2.2 (Schubert *et al.*, 2016), and the quality of the reads was checked with FastQC version 0.11.5 (Andrews, 2010). Reports were combined using MultiQC version 1.0 (Ewels *et al.*, 2016). Open reading frames (ORFs) were predicted on the unmerged and

unassembled forward trimmed reads using FragGeneScan-Plus (Kim *et al*., 2015). Profile hidden Markov models (HMMs) for taxonomic marker (*rpoB*) and functional genes (*amoA*_AOA, *amoA*_AOB, and *nxrB*), downloaded from FunGene (Fish *et al.*, 2013), were used to quantify the relative abundances and taxonomic affiliations of nitrifiers from the unassembled reads using MetAnnotate (Petrenko *et al.*, 2015). The HMM e-value threshold used by MetAnnotate for gene detection was 10^{-3}, and the e-value threshold used for the USEARCH-based taxonomic classification step was 10^{-6}. Separate *amoA* HMMs for AOA and AOB were used for improved gene recovery. Within MetAnnotate, the database used for taxonomic classification was RefSeq release 80 (March 2017). MetAnnotate results were analyzed using the custom R script *metannotate_barplots.R* version 0.9 (available at https://github.com/jmtsuji/metannotate-analysis). To allow for approximate between-sample and between-HMM comparisons, hit counts for each HMM were normalized both to the HMM length and to the total number of length-normalized HMM hits for *rpoB* (for more details, see the Github README).

Assembly and binning of metagenomic reads

Raw metagenomic reads were additionally processed through the ATLAS pipeline (release 1.0.22), which includes quality control, assembly, annotation, binning, and read mapping of metagenomic data (White *et al*., 2017). For further details, see supplemental methods. Bin completeness and contamination were assessed using CheckM version 1.0.7 (Parks *et al*., 2015). *Nitrospira* genome bins from assemblies were compared and dereplicated to find identical genomes using dRep version 2.0.5 (Olm *et al.*, 2017). Average nucleotide identity (ANI) analysis was performed on the representative ("winning") AOA genome selected by dRep, as well as the high quality (>75% completion, <25% contamination) *amoA*-containing *Nitrospira* bins, using

ANIm on the JSpecies webserver (Richter *et al.*, 2016). Sequence data was deposited in the European Nucleotide Archive under study accession number PRJEB30654.

Phylogenetic analysis

All *amoA* genes from assembled contigs that were classified as the genus *Nitrospira* were compiled and clustered into groups based on nucleotide sequence at a 99% identity threshold using CD-HIT version 4.7 (Li and Godzik, 2006). Using MUSCLE (Edgar, 2004) within MEGA7 (Kumar *et al.*, 2016), the corresponding amino acid sequences of these genes were aligned, along with comammox AmoA sequences from cultivated species and environmental surveys retrieved from the NCBI database. Phylogenetic analysis was inferred using maximum likelihood analysis with 500 bootstrap replicates, based on the Le Gascuel evolutionary model (Le and Gascuel, 2008), using MEGA7.

Abundance of comammox bacteria bins

Representative bins for each *amoA* RBC cluster were selected based on dRep results. Read mapping (using quality trimmed reads) to representative bins was performed using BBMap (Bushnell, 2017) version 38.22. Reads from all samples were iteratively mapped to each bin, and average depth of coverage per nucleotide was calculated. For further details, see supplemental methods.

Results

Wastewater nitrogen concentrations

Ammonium concentrations were higher in RBC 1 (above 9 μM) than RBC 8 (below 4 μM) for all treatment trains (Figure 1C, Table S4). Nitrite concentrations were higher than ammonium concentrations for most sampled RBCs, and nitrate concentrations greatly exceeded ammonia and nitrite (Table S4). The average plant influent and effluent ammonia concentrations were relatively stable over the years 2010 to 2016 (Figure S1A). As an assessment of the long-term temporal stability of ammonia treatment in this WWTP, plant data demonstrate that ammonia concentrations of the RBC influent were lowest in 2012 (16.4 μM) and 2016 (22.1 μM) (Figure S1B). The average ammonia influent concentration in 2016 was lower than in 2011 (81.4 μM); however, this was not a significant difference (t-test, $p = 0.07$).

Gene abundances in biofilm samples

Thaumarchaeotal 16S rRNA genes, AOB *amoA*, and comammox *amoA* genes were detected in all samples via qPCR (Figure 2, Figure S2). The number of comammox *amoA* genes, as well as the proportion of comammox bacteria to the total community was consistently higher in RBC 1 than RBC 8 across all sample pairs (Figure 2, Table S5). Thaumarchaeotal 16S rRNA genes were most abundant in RBC 8 and less abundant in RBC 1 for 2010 samples, and the proportion of AOA to the total community was consistently higher in RBC 8 than RBC 1 for the same month in 2010. In contrast, AOA 16S rRNA gene abundances were higher in RBC 1 than RBC 8 in all four trains sampled in 2016 and roughly equivalent in abundance to comammox bacteria (Figure 2). The AOB *amoA* gene abundances were higher in RBC 1 than RBC 8 in both years, but in 2016 their abundances were one to two orders of magnitude lower than in 2010.

In most samples, comammox *amoA* genes were more abundant than those of other ammonia oxidizers (Figure 2, Figure S2, Table S5). In 2010, comammox *amoA* outnumbered AOA 16S rRNA and AOB *amoA* genes in all samples, except for the September RBC 8 sample, in which AOA genes outnumbered the genes of the other ammonia oxidizers. In the 2010 samples, AOB *amoA* genes outnumbered AOA 16S rRNA genes in RBC 1 at all timepoints, whereas AOA genes either outnumbered AOB genes in RBC 8, or their abundances were roughly equal. In 2016, comammox *amoA* genes outnumbered the genes of the other ammonia oxidizers, except in NE8 and NW8, in which AOA genes outnumbered comammox genes. The AOA genes greatly outnumbered AOB genes in RBC 1 and 8 in all trains in the 2016 samples. Analysis of all 2016 RBC stages showed a general trend of decreasing *amoA* gene abundance across each RBC flowpath (Figure S2).

Biofilm community composition from metagenomics

Based on taxonomic marker gene analysis of unassembled metagenomic sequence data, the *Nitrospira* sequences dominated all samples and comprised between 8 and 32% of the total *rpoB* sequences among RBC samples from the two sampling years (Figure 3). The *Nitrospira*-affiliated *rpoB* gene relative abundances were higher in RBC 1 than the corresponding RBC 8 for each train sampled. Consistent with qPCR data, *Nitrosomonas*-affiliated *rpoB* gene sequences were detected at ≥1% in the 2010 samples. The community composition of the dominant microorganisms was stable between seasonal timepoints in 2010, and these profiles were different than those detected in 2016. The main difference was the decrease in *Nitrosomonas rpoB* gene abundances in 2016, making *Nitrospira* the only potential ammonia oxidizer present

at ≥1% abundance based on this analysis. The *rpoB* gene of AOA was not detected at ≥1% relative abundance in any of the samples.

The HMM search for the *amoA* gene of AOA detected this gene at low relative abundance in the metagenomic dataset (<2% relative to *rpoB*; Figure 4A). For 2010 samples, *amoA* gene sequences affiliated with *Thaumarchaeota* were more abundant in RBC 8 than RBC 1 samples, whereas in 2016 they were more abundant in RBC 1 than RBC 8 samples, except in the NE 2016 sample. The HMM hits for AOA *amoA* genes in the unassembled metagenomes were low, with only 64 hits to this HMM for all samples combined. Despite this, the pattern of *amoA* sequence abundance (normalized to *rpoB*) correlated significantly with the corresponding relative abundances determined via qPCR for the proportion of AOA 16S rRNA genes within the total community (Spearman's rank correlation, $r_s = 0.67$, $p = 0.009$).

The *amoA* HMM for AOB detected sequences of both potential comammox bacteria and AOB. The *amoA* sequence hits that were taxonomically assigned to *Nitrospira* are indicative of comammox bacteria. Overall, the relative abundances of AOB *amoA* genes were higher in RBC 1 than RBC 8 of the corresponding sample pairs, in both sampling years (Figure 4B). In 2010, *Nitrosomonas* spp. and *Nitrospira* spp. were at roughly equal relative abundances, whereas *Nitrospira* spp. were the dominant ammonia oxidizers detected in 2016 metagenomes. The *rpoB* and *Nitrospira*-associated *amoA* gene abundances were approximately equal (Figure 3, Figure 4B), suggesting that a substantial proportion of the *Nitrospira* present in the RBCs were comammox bacteria. Across all samples, comammox bacteria represented between ~5-30% of the total microbial community based on comparing *amoA* HMM hits to total *rpoB* HMM hits, assuming that each comammox cell encodes one copy of the *amoA* gene. These relative abundance patterns for AOB and comammox bacteria *amoA* genes were significantly correlated

to their corresponding relative abundances determined via qPCR for the proportion of their *amoA* genes within the total community (AOB $r_s = 0.97$, $p < 0.001$; comammox $r_s = 0.79$, $p < 0.001$).

The HMM search for *nxrB* revealed that sequences affiliated with this model were highly prevalent in the RBC biofilm metagenomes, with combined hits at over 50% abundance relative to total *rpoB* hits (Figure 4C). This indicates a high relative abundance of microorganisms that could potentially oxidize nitrite. The high relative abundance of *Nitrospira*-affiliated *nxrB* gene sequences, which were more abundant in RBC 1 than RBC 8 metagenomes across all samples, is further evidence for *Nitrospira* dominance in the RBCs. The relative abundances of *Nitrospira*-affiliated *nxrB* sequences were likely higher than for *rpoB* sequences because most *Nitrospira* spp. have multiple copies of *nxrB* (Pester *et al.*, 2014), whereas *rpoB* is typically a single copy gene (Dahllof *et al.*, 2000).

Recovery of genome bins from metagenomes

Both individual and co-assembly methods were used (Table S6) to recover genome bins from metagenomic data, resulting in a collection of 53 high quality (≥75% completeness, <25% contamination) genome bins classified as *Nitrospira*. Clustering and dereplication reduced this to a collection of 16 genome bins representing unique *Nitrospira* populations in the RBCs (Table 1, Figure S3). Seven high quality *Nitrospira* bins contained *amoA* genes and were dereplicated into five different clusters. Distinct clades of *amoA*-containing bins were observed, as well as a clade with no *amoA*-containing bins, indicating that the *Nitrospira* community in the RBCs was composed of both comammox *Nitrospira* and strict nitrite-oxidizing *Nitrospira*. Using average nucleotide identity (ANI) analysis, the seven high-quality *Nitrospira* bins containing *amoA* were compared to each other, to cultivated comammox bacterial genomes, and to genomes of strict

nitrite-oxidizing *Nitrospira* (Table S7). Bins with >99% ANI values belonged to the same primary dRep cluster, indicating the robustness of the dereplication process. All of these comammox bins had less than 95% ANI to any of the cultivated comammox bacterial genomes or the NOB genomes, suggesting that these bins represent distinct comammox species from cultivated representatives.

Evidence of complete ammonia oxidation metabolism in the *Nitrospira* of the RBCs is demonstrated by the combined presence of *amo, hao,* and *nxr* genes which are known to be encoded by comammox bacteria (van Kessel *et al.*, 2015; Daims *et al.*, 2015). Two bins from the individual assembly attempts, NE8.001 and SNE8.002, contained *amoAB, hao,* and *nxrB* genes (Table 1). These bins were nearly complete (92.2%) with low contamination (2.7%), and substantially complete (76.8%) with medium contamination (5.8%), respectively. These bins also represent distinct comammox populations as identified by dRep (Figure S3). The comammox bacteria identified in this study likely have the ability to hydrolyze urea, as indicated by a nearly complete *ure* operon on contigs belonging to *amoA*-containing *Nitrospira* bins (Table 1). These *Nitrospira* bins did not contain genes for formate dehydrogenase (*fdh*).

Co-assembly and differential abundance binning was also used to obtain *Thaumarchaeota* bins. Dereplicated *Thaumarchaeota* genome bins were all >99% similar (Supplementary file 1), indicating the presence of a single *Thaumarchaeota* species. Detection of a single species matches with previous data from the RBCs (Sauder *et al.*, 2012, 2017). An average nucleotide identity (ANI) of 96.8% between the dereplicated AOA bin (CA-NE8-2010.39) and the *Ca.* Nitrosocosmicus exaquare genome (GenBank accession number

CP017922.1) indicated that the AOA genome bin corresponded to the same archaeon previously enriched from RBC biofilm samples.

Phylogenetic relationships

From all assembled contigs, a total of 71 *amoA* gene sequences were classified as *Nitrospira*; of these, 19 were assigned to genome bins (Table S8). Bin SW8.001 contained two *amoA* genes, whereas the rest had only one *amoA* gene copy. All sequenced *amoA* genes of *Nitrospira* spp. belonged to comammox clade A (Figure 5). There were as many as 13 distinct groups of comammox *Nitrospira* based on clustering of the *amoA* gene by sequence identity (Figure 5, Table S8), although it is possible that the number of groups could be lower if closely related *amoA* gene sequences represent the same population.

A majority of the *amoA* gene sequences (i.e., RBC groups B, C, D, E, F, I, L, and M) clustered phylogenetically with uncultivated comammox species from a metagenomic survey of water treatment systems (Wang *et al.*, 2017) and a rice paddy soil from China (Genbank accession AKD44274). The other *amoA* gene sequence groups clustered more closely to cultivated species of comammox bacteria or uncultivated comammox species detected from other environments, including drinking water systems and another WWTP (Figure 5).

Average coverage was used as a measure of abundance for each bin across the samples (Figure 5). Bins NW1.001 and SNE1.002, which contained *amoA* genes corresponding to the RBC Groups F and B, respectively, showed similar abundance distributions across the RBC samples. Bin NE8.001, containing the *amoA* gene corresponding to RBC Group E, also showed similar abundances to these bins. All three of these bins represent closely related RBC groups, according to AmoA phylogeny (Figure 5). Across all samples, these bins also appeared to be the

most abundant (as measured by average bin coverage), and NE8.001 represented the most abundant population of comammox in 2016. At the same time, all genome bins were found to have a small proportion of shared reads even when read mapping settings required exact sequence matches (data not shown), implying some degree of shared genome sequence information. As one extreme case, one pair of dereplicated *Nitrospira* genome bins, SNE8.002 (Group B; Figure 5) and NW1.001 (Group F), shared as much as 20% of their total mapped reads depending on the metagenome sample being mapped (data not shown). Bins SNE8.002 and NW1.001 were classified via dRep into the same primary cluster but were split into different secondary clusters (Figure S3), implying that they could either represent the same population or closely related strains. Because of this, read mapping data was not used to quantitatively track co-occurrence of comammox genome bins but was used to make qualitative observations about the general distribution of recovered bins throughout the RBC samples. Multiple comammox populations were found in the same RBC. This high diversity of comammox bacteria contrasts the low AOA diversity (i.e., only one species) in the RBCs.

Discussion

Both qPCR and metagenomic data confirmed that *amoA*-containing *Nitrospira* (i.e., comammox bacteria) were numerically dominant ammonia oxidizers in the RBC biofilm. Due to their recent discovery (Daims *et al.*, 2015; van Kessel *et al.*, 2015), comammox bacteria were not previously investigated at the Guelph WWTP, and previously analyses likely failed to detect the most abundant ammonia oxidizers present (Sauder *et al.*, 2017, 2012). Not only were comammox bacteria the dominant ammonia oxidizers based on qPCR (Figure 2), they

represented the dominant microbial community members based on taxonomic assignment of metagenomic sequences (Figure 3, Figure 4). These results contrast with previous studies of full-scale activated sludge WWTPs that have reported comammox bacteria at lower relative abundances than AOA and other AOB (Pjevac *et al.*, 2017; Chao *et al.*, 2016; Fan *et al.*, 2017; Pan *et al.*, 2018).

At the Guelph WWTP, the RBCs are located downstream of the activated sludge aeration basins. Ammonia concentrations entering this tertiary treatment system are approximately two orders of magnitude lower than within aeration basins. For example, on the day of sampling in 2016, the ammonia concentration in aeration basin influents averaged 2.5 mM NH_3-N (24 hour composite sample, data obtained from WWTP operators). This is typical of aerobic secondary aeration basins (e.g. Fan *et al.*, 2017; Wang *et al.*, 2018; Pan *et al.*, 2018; Park *et al.*, 2006). The predicted low ammonia niche for comammox bacteria and their biofilm growth (Costa *et al.*, 2006; Daims *et al.*, 2015) may explain the dominance of comammox bacteria in this tertiary treatment system despite the reported low abundances in secondary activated sludge flocs (Pjevac *et al.*, 2017; Chao *et al.*, 2016; Fan *et al.*, 2017). However, ammonia concentrations do not explain all comammox bacterial abundance patterns. In contrast to these other published studies, a recent PCR-based study found that comammox *amoA* genes were more abundant than those of AOB in activated sludge samples from eight separate WWTPs, despite a high ammonia-microorganism ratio (Wang *et al.*, 2018). This suggests that, as for AOA (e.g., Bai *et al.*, 2012; Roy *et al.*, 2017; Gao *et al.*, 2013; Zhang *et al.*, 2011; Limpiyakorn *et al.*, 2011), the abundances of comammox bacteria vary across WWTPs, and more studies are needed to determine the factors in WWTP environments that affect comammox abundances. The current discovery of dominant comammox bacterial communities in the Guelph RBCs suggests a correspondingly

important role for nitrification activity by comammox bacteria in a full scale wastewater treatment application.

High abundances of comammox bacteria have been reported for other engineered environments with low ammonia levels, such as drinking water systems. Comammox bacteria have been found in a drinking water treatment plant (Pinto *et al.*, 2015) and distribution systems (Wang *et al.*, 2017), and they were more abundant than other ammonia oxidizers in a groundwater well studied by Pjevac *et al.* (2017). In drinking water treatment plants, comammox bacteria were found to be the most abundant nitrifier in groundwater fed rapid sand filters (Fowler *et al.*, 2018). In those filters, clade B comammox bacteria were the dominant group, in contrast to the clade A comammox bacteria detected in the sampled RBCs. A study examining recirculating aquaculture system biofilters found that comammox bacteria were more abundant than AOA and AOB (Bartleme *et al.*, 2017). Along with these other studies, the dominance of comammox bacteria in the RBCs suggests comammox bacteria thrive in engineered environments with relatively low ammonia concentrations.

The high number of unique *Nitrospira amoA* genes recovered from the RBC metagenomes indicates high diversity of comammox bacteria within a single WWTP (Figure 5). Thirteen unique *Nitrospira amoA* genes were recovered at 99% sequence clustering, and although only seven could be placed into high-quality genome bins (Figure S3, Table S8), it is likely that many of the unbinned *Nitrospira amoA* genes also represent unique comammox populations given the low copy number per genome typical for *amoA* (Daims *et al.*, 2015; van Kessel *et al.*, 2015). Multiple populations of comammox bacteria within the same WWTP have been reported previously (Pjevac *et al.*, 2017; Wang *et al.*, 2018), but with lower observed diversity (i.e. one to eight operational taxonomic units) than reported here. The diversity of

potential comammox bacteria in the RBCs also appears to exceed that of AOA and AOB in the RBCs, which had fewer genome bins identified (Supplementary file 1). In fact, some of the apparent *Nitrospira* NOB genome bins (i.e., lacking an ammonia oxidation gene system; Table 1) may represent comammox bins that are missing the *amoA* gene due to incomplete binning. The presence of many populations of comammox bacteria in the same RBC suggests that comammox bacteria may not compete directly, but may occupy unique niches within the RBCs, for example because of distinct microenvironments within the biofilm that vary in oxygen or ammonia concentrations. The abundances of the comammox bins varied over time and between RBCs (Figure 5), implying that complex physicochemical factors govern comammox distribution in the RBCs. Some comammox genome bins may even represent closely related strains (Results; Figure S3), adding further interest to understanding the factors governing the distribution and ecology of these organisms. A similar high diversity of comammox bacteria may also exist in other WWTP systems.

All detected *Nitrospira amoA* gene sequences grouped phylogenetically in comammox bacterial clade A (Figure 5), consistent with other studies reporting comammox bacteria in WWTPs (Pjevac *et al.*, 2017; Chao *et al.*, 2016). Recently, Xia *et al.* (2018) proposed that clade A could be further subdivided into clade A.1, which includes the cultivated comammox *Nitrospira*, and clade A.2, which contains only uncultivated comammox AmoA sequences, including those from drinking water systems (OQW38018 and OQW37964; Wang *et al.*, 2017). *Nitrospira amoA* gene sequences detected in the RBCs were classified into both of these clade A subdivisions (Figure 5). RBC Groups A, G, H, K, and J appear to fall into clade A.1, grouping closely with cultivated species of comammox bacteria, while the other RBC Groups (I, E, C, D, F, B, M, and L) appear to fall into the newly proposed clade A.2 with no cultivated

representative. Sequences of clade A.2 AmoA from the RBCs were most closely related to AmoA sequences obtained from metagenomes of drinking water systems (Wang *et al.*, 2017) or to a clone library sequence from a rice paddy (Genbank accession AKD44274). Comammox bacteria related to the rice paddy clone have been detected in full-scale WWTPs, and this clade was also described as a new comammox clade A cluster by the authors (Wang *et al.*, 2018). Comparisons of the high quality comammox bacterial bins to cultivated comammox bacterial genomes show low ANI (Table S7), which is further evidence that at least some of the comammox bacteria in the RBCs represent novel species.

Although the presence of *amoA* is an indication the *Nitrospira* are likely comammox bacteria, further evidence of complete ammonia oxidation metabolism in the *Nitrospira* of the RBCs is demonstrated by finding genes for both ammonia oxidation and nitrite oxidation in the same high quality genome bin (Table, Figure S3). *Nitrospira* in the RBC biofilm have previously been demonstrated to incorporate labelled bicarbonate in the presence of ammonia (Sauder *et al.*, 2017), indicating autotrophic activity of the organisms upon addition of inorganic nitrogen. The presence of a *ure* operon (Table 1) suggests that these comammox bacteria of the RBCs are capable of hydrolyzing urea to ammonia, further supporting an ammonia-based metabolism. The presence of these genes is consistent with clade A comammox bacteria and comammox enrichment cultures that utilize urea (van Kessel *et al.*, 2015; Daims *et al.*, 2015; Palomo *et al.*, 2018; Lawson and Lücker, 2018). Additionally, the sequenced RBC comammox bacteria likely cannot use formate as an alternative electron donor, unlike NOB *Nitrospira*, because the *amoA*-containing *Nitrospira* bins did not contain any *fdh* genes for formate dehydrogenase (Table 1), which is consistent with the findings of Palomo *et al.* (2018).

Consistent and distinct abundance patterns of taxa in RBC 1 versus RBC 8 in respective sampling years (Figure 2, Figure 3, Figure 4) imply a relationship between the environmental conditions of the RBCs and the microbial ammonia-oxidizing community. Ammonia concentrations are a likely factor affecting ammonia oxidizer distributions. Comammox bacteria were predicted to have higher abundances in RBC 8 than RBC 1 due to the predicted low ammonia niche of comammox bacteria (Costa *et al.*, 2006), but for both years tested, the opposite trend was observed (Figure 2, Figure 4B, Figure S2). However, the ammonia concentrations in the first RBCs were already low (<18 μM for the 2016 samples, Figure 1C), and it may be that the ammonia concentrations in RBC 8 influent (<4 μM) were below the optimum for these comammox species. The ammonia affinity of only one comammox species (*Ca.* Nitrospira inopinata) has been determined, and it has one of the highest reported affinities of all ammonia oxidizers (reported $K_{m(app)}$ of 0.65 to 1.1 μM total ammonium) (Kits *et al.*, 2017). The affinities of the RBC comammox bacteria are unknown, but given their higher abundance in RBC 1 than RBC 8, they may prefer higher ammonia concentrations than *Ca.* N. inopinata. Additionally, a smaller population of nitrifiers may be supported in RBC 8 than RBC 1 because there is less total ammonia available. Other factors that likely also lead to the differences in relative abundances of comammox bacteria across RBCs could include oxygen requirements, growth rates, growth yields, and biofilm formation capabilities (Lawson and Lücker, 2018).

The AOA relative abundances measured by qPCR (Figure 2) and metagenomic analyses (Figure 4A) differed, such that AOA were underrepresented in metagenomes. Previous work on the Guelph RBCs detected similar abundances of AOA in RBC samples with both qPCR and quantitative fluorescent *in situ* hybridization (FISH) (Sauder *et al.*, 2017), which suggests that the qPCR data provide a more representative estimate of AOA numbers within the sampled

biofilm. Future research would be required to assess whether low GC content of the *Ca.* N. exaquare genome (33.9%) or HMM bias against AOA sequences may contribute to underrepresentation of AOA within metagenomes prepared by high-throughput sequencing on the Illumina platform. Nonetheless, both qPCR and metagenomics suggested the same overall spatial and temporal distribution patterns for AOA within the RBCs.

The observed pattern of AOA *amoA* abundances for 2010 samples analyzed in this study (Figure 2, Figure 4A) was consistent with previous data (Sauder *et al.*, 2012), specifically in that AOA were predicted to be at higher relative abundances in RBC 8 than RBC 1. Although AOA in 2016 samples were at a lower relative abundance in RBC 8 compared to RBC 1, this matches data from more recent samples collected from the RBCs in December 2015 (Sauder *et al.*, 2017). We speculate that the reversal in the AOA relative abundance pattern was caused by lower overall ammonia concentrations that have been measured in the Guelph tertiary treatment system in recent years (Figure S1B) and by cleaning of the RBC trains in 2014 that would have disturbed the microbial communities. These changes to the RBCs may have impacted the AOB as well. Both qPCR (Figure 2) and metagenomic data (Figure 4B) indicated an overall decrease in AOB *amoA* gene relative abundances in the 2016 samples compared to 2010. Most AOB have a low ammonia affinity (Lehtovirta-Morley, 2018; Kits *et al.*, 2017; Prosser and Nicol, 2012), and therefore the lower ammonia levels entering the RBCs in 2016 might have led to AOB being outcompeted by ammonia oxidizers with a higher ammonia affinity. The relative abundances of comammox *Nitrospira*, meanwhile, were more consistent temporally and spatially than those of the AOA and AOB at the genus level (Figure 4), although relative abundances were quite variable across all samples at the genome level (Figure 5). This suggests that the factors that

drive AOA, AOB, and comammox abundance and diversity are distinct within the RBC biofilm environment.

This study is among the first to report comammox bacteria as the dominant ammonia oxidizer in an engineered WWTP system. Detected comammox bacteria exhibited a high level of overall diversity, and higher diversity than other ammonia oxidizers. Although comammox bacteria were numerically dominant, their contributions to nitrification in the RBC environment remains unclear. Future studies using differential inhibitors and laboratory cultivation may be useful for elucidating the contributions of comammox bacteria to nitrification in the RBC biofilms. In addition, the comammox that fall within the newly proposed clade A.2 would be ideal targets for cultivation given that no cultured representative is yet available for this group.

The Guelph WWTP RBCs represent a unique and useful system to study the ecology of comammox bacteria as a result of built-in environmental gradients and the high relative abundances of comammox bacteria, AOA, and traditional AOB. Tertiary treatment systems such as these RBCs can be used to produce high quality effluents, which may be particularly important when WWTPs discharge into environments that have low assimilative capacities or are ecologically sensitive. Understanding the microorganisms that are present and their contributions to nitrification within the context of water treatment is an important step towards understanding the microbial contributions to nitrification in low-ammonia biofilm systems and may contribute to knowledge of operational practices for improved effluent quality in municipal, industrial, and aquaculture-associated water treatment systems.

Acknowledgements

We thank the operators and staff at the Guelph WWTP for access to their facility and for providing plant data. Funding for metagenomic sequencing was provided by the Faculties of Science and Engineering at the University of Waterloo. This research was supported by Discovery Grants to LAH, ACD, WJP, and JDN from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Competing interests

The authors declare no competing interests.

References

Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Bai Y, Sun Q, Wen D, Tang X. (2012). Abundance of ammonia-oxidizing bacteria and archaea in industrial and domestic wastewater treatment systems. *FEMS Microbiol Ecol* **80**: 323–330.

Bartleme RP, Mclellan SL, Newton RJ. (2017). Freshwater recirculating aquaculture system operations drive biofilter bacterial community shifts around a stable nitrifying consortium of ammonia-oxidizing archaea and comammox *Nitrospira*. *Front Microbiol* **8**: 101.

Bushnell B. (2017). BBMap. https://sourceforge.net/projects/bbmap/.

Chao Y, Mao Y, Yu K, Zhang T. (2016). Novel nitrifiers and comammox in a full-scale hybrid biofilm and activated sludge reactor revealed by metagenomic approach. *Appl Microbiol Biotechnol* **100**: 8225–8237.

Costa E, Pérez J, Kreft J-U. (2006). Why is metabolic labour divided in nitrification? *Trends Microbiol* **14**: 213–219.

Dahllof I, Baillie H, Kjelleberg S. (2000). *rpoB*-based microbial community analysis avoids limitations inherent in 16S rRNA gene intraspecies heterogeneity. *Appl Environ Microbiol* **66**: 3376–3380.

Daims H, Lebedeva E V., Pjevac P, Han P, Herbold C, Albertsen M, *et al.* (2015). Complete nitrification by *Nitrospira* bacteria. *Nature* **528**: 504–509.

Daims H, Lücker S, Wagner M. (2016). A new perspective on microbes formerly known as nitrite-oxidizing bacteria. *Trends Microbiol* **24**: 699–712.

Edgar RC. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.

Ewels P, Magnusson M, Lundin S, Käller M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**: 3047–3048.

Fan X-Y, Gao J-F, Pan K-L, Li D-C, Dai H-H. (2017). Temporal dynamics of bacterial communities and predicted nitrogen metabolism genes in a full-scale wastewater treatment plant. *RSC Adv* **7**: 56317–56327.

Fish JA, Chai B, Wang Q, Sun Y, Brown CT, Tiedje JM, *et al.* (2013). FunGene: the functional gene pipeline and repository. *Front Microbiol* **4**: 291.

Fowler SJ, Palomo A, Dechesne A, Mines PD, Smets BF. (2018). Comammox *Nitrospira* are abundant ammonia oxidizers in diverse groundwater-fed rapid sand filter communities. *Environ Microbiol* **20**: 1002–1015.

Gao J-F, Luo X, Wu G-X, Li T, Peng Y-Z. (2013). Quantitative analyses of the composition and abundance of ammonia-oxidizing archaea and ammonia-oxidizing bacteria in eight full-scale biological wastewater treatment plants. *Bioresour Technol* **138**: 285–296.

Gonzalez-Martinez A, Rodriguez-Sanchez A, van Loosdrecht MCM, Gonzalez-Lopez J, Vahala R. (2016). Detection of comammox bacteria in full-scale wastewater treatment bioreactors using tag-454-pyrosequencing. *Environ Sci Pollut Res* **23**: 25501–25511.

Holmes RM, Aminot A, Kérouel R, Hooker BA, Peterson BJ. (1999). A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Can J Fish Aquat Sci* **56**: 1801–1808.

van Kessel MAHJ, Speth DR, Albertsen M, Nielsen PH, Op den Camp HJM, Kartal B, *et al.* (2015). Complete nitrification by a single microorganism. *Nature* **528**: 555–559.

Kim D, Hahn AS, Wu SJ, Hanson NW, Konwar KM, Hallam SJ. (2015). FragGeneScan-plus for scalable high-throughput short-read open reading frame prediction. In: *2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2015*. IEEE, pp 1–8.

Kits KD, Sedlacek CJ, Lebedeva E V., Han P, Bulaev A, Pjevac P, *et al.* (2017). Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature* **549**: 269–272.

Könneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543–546.

Kumar S, Stecher G, Tamura K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* **33**: 1870–1874.

Lawson CE, Lücker S. (2018). Complete ammonia oxidation: an important control on nitrification in engineered ecosystems? *Curr Opin Biotechnol* **50**: 158–165.

Le SQ, Gascuel O. (2008). An improved general amino acid replacement matrix. *Mol Biol Evol* **25**: 1307–1320.

Lehtovirta-Morley LE. (2018). Ammonia oxidation: Ecology, physiology, biochemistry and why they must all come together. *FEMS Microbiol Lett* **365**: fny058.

Li W, Godzik A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**: 1658–1659.

Limpiyakorn T, Sonthiphand P, Rongsayamanont C, Polprasert C. (2011). Abundance of *amoA* genes of ammonia-oxidizing archaea and bacteria in activated sludge of full-scale wastewater treatment plants. *Bioresour Technol* **102**: 3694–3701.

Miranda KM, Espey MG, Wink DA. (2001). A rapid, simple spectrophotometric method for simultaneous detection of nitrate and nitrite. *Nitric Oxide* **5**: 62–71.

Muyzer G, de Waal EC, Uitterlinden AG. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695–700.

Ochsenreiter T, Selezi D, Quaiser A, Bonch-Osmolovskaya L, Schleper C. (2003). Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR. *Environ Microbiol* **5**: 787–797.

Olm MR, Brown CT, Brooks B, Banfield JF. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication.

ISME J **11**: 2864–2868.

Palomo A, Fowler SJ, Gülay A, Rasmussen S, Sicheritz-Ponten T, Smets BF. (2016). Metagenomic analysis of rapid gravity sand filter microbial communities suggests novel physiology of *Nitrospira* spp. *ISME J* **10**: 2569–2581.

Palomo A, Pedersen AG, Fowler SJ, Dechesne A, Sicheritz-Pontén T, Smets BF. (2018). Comparative genomics sheds light on niche differentiation and the evolutionary history of comammox *Nitrospira*. *ISME J* **12**: 1779–1793.

Pan KL, Gao JF, Fan XY, Li DC, Dai HH. (2018). The more important role of archaea than bacteria in nitrification of wastewater treatment plants in cold season despite their numerical relationships. *Water Res* **145**: 552–561.

Park H-D, Wells GF, Bae H, Criddle CS, Francis CA. (2006). Occurrence of ammonia-oxidizing archaea in wastewater treatment plant bioreactors. *Appl Environ Microbiol* **72**: 5643–5647.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.

Pester M, Maixner F, Berry D, Rattei T, Koch H, Lücker S, *et al*. (2014). *NxrB* encoding the beta subunit of nitrite oxidoreductase as functional and phylogenetic marker for nitrite-oxidizing *Nitrospira*. *Environ Microbiol* **16**: 3055–3071.

Petrenko P, Lobb B, Kurtz DA, Neufeld JD, Doxey AC. (2015). MetAnnotate: function-specific taxonomic profiling and comparison of metagenomes. *BMC Biol* **13**: 92.

Pinto AJ, Marcus DN, Ijaz Z, Bautista-de los Santos QM, Dick GJ, Raskin L. (2015).

Metagenomic evidence for the presence of comammox *Nitrospira*-like bacteria in a drinking water system. *mSphere* **1**: e00054-15.

Pjevac P, Schauberger C, Poghosyan L, Herbold CW, Kessel MA van, Daebeler A, *et al.* (2017). *AmoA*-targeted polymerase chain reaction primers for the specific detection and quantification of comammox *Nitrospira* in the environment. *Front Microbiol* **8**: 1508.

Poulin P, Pelletier É. (2007). Determination of ammonium using a microplate-based fluorometric technique. *Talanta* **71**: 1500–1506.

Prosser JI, Nicol GW. (2012). Archaeal and bacterial ammonia-oxidisers in soil: The quest for niche specialisation and differentiation. *Trends Microbiol* **20**: 523–531.

Purkhold U, Pommerening-Röser A, Juretschko S, Schmid MC, Koops HP, Wagner M. (2000). Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA and *amoA* sequence analysis: Implications for molecular diversity surveys. *Appl Environ Microbiol* **66**: 5368–5382.

Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. (2016). JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* **32**: 929–931.

Rotthauwe JH, Witzel KP, Liesack W. (1997). The ammonia monooxygenase structural gene *amoA* as a functional marker: Molecular fine-scale analysis of natural ammonia-oxidizing populations. *Appl Environ Microbiol* **63**: 4704–4712.

Roy D, McEvoy J, Blonigen M, Amundson M, Khan E. (2017). Seasonal variation and ex-situ nitrification activity of ammonia oxidizing archaea in biofilm based wastewater treatment

processes. *Bioresour Technol* **244**: 850–859.

Sauder LA, Albertsen M, Engel K, Schwarz J, Nielsen PH, Wagner M, *et al.* (2017). Cultivation and characterization of *Candidatus* Nitrosocosmicus exaquare, an ammonia-oxidizing archaeon from a municipal wastewater treatment system. *ISME J* **11**: 1142–1157.

Sauder LA, Peterse F, Schouten S, Neufeld JD. (2012). Low-ammonia niche of ammonia-oxidizing archaea in rotating biological contactors of a municipal wastewater treatment plant. *Environ Microbiol* **14**: 2589–2600.

Schubert M, Lindgreen S, Orlando L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* **9**: 88.

Stein LY, Klotz MG. (2016). The nitrogen cycle. *Curr Biol* **26**: R94–R98.

Wang M, Huang G, Zhao Z, Dang C, Liu W, Zheng M. (2018). Newly designed primer pair revealed dominant and diverse comammox *amoA* gene in full-scale wastewater treatment plants. *Bioresour Technol* **270**: 580–587.

Wang Y, Ma L, Mao Y, Jiang X, Xia Y, Yu K, *et al.* (2017). Comammox in drinking water systems. *Water Res* **116**: 332–341.

White RAI, Brown J, Colby S, Overall CC, Lee J-Y, Zucker J, *et al.* (2017). ATLAS (Automatic Tool for Local Assembly Structures) -a comprehensive infrastructure for assembly, annotation, and genomic binning of metagenomic and metatranscriptomic data. *PeerJ Prepr* 1–11.

Xia F, Wang J-G, Zhu T, Zou B, Rhee S-K, Quan Z-X. (2018). Ubiquity and diversity of complete ammonia oxidizers (comammox). *Appl Environ Microbiol* **84**: e01390-18.

Zhang T, Ye L, Tong AHY, Shao MF, Lok S. (2011). Ammonia-oxidizing archaea and

ammonia-oxidizing bacteria in six full-scale wastewater treatment bioreactors. *Appl Microbiol Biotechnol* **91**: 1215–1225.

Zhao Z, Huang G, Wang M, Zhou N, He S, Dang C, *et al*. (2018). Abundance and community composition of comammox bacteria in different ecosystems by a universal primer set. *bioRxiv* 492488.

Zheng M, Wang M, Zhao Z, Zhou N, He S, Liu S, *et al.* (2019). Transcriptional activity and diversity of comammox bacteria as a previously overlooked ammonia oxidizing prokaryote in full-scale wastewater treatment plants. *Sci Total Environ* **656**: 717–722.

Figure legends

Figure 1 Rotating biological contactors (RBCs) of the Guelph wastewater treatment plant (WWTP). (A) External view of the RBCs and (B) schematic of the RBC trains. Arrows indicate the direction of water flow. (C) Ammonium concentrations (NH_4^+-N) in RBC influent water sampled at the same time as biofilm collection in October 2016. Error bars indicate standard deviation of technical duplicates. NE, Northeast train; NW, Northwest train; SE, Southeast train; SW, Southwest train.

Figure 2 Bacterial 16S rRNA, comammox and AOB *amoA*, and archaeal 16S rRNA gene abundances for samples paired with metagenome sequencing. Error bars indicate standard deviation of technical qPCR duplicates. Pie charts show gene abundances as a proportion of all ammonia oxidizing prokaryotes. Numbers above the pie charts indicate the proportion (%) of the total community that all ammonia oxidizing prokaryotes represent. NE, northeast train; NW, northwest train; SE, southeast train; SW, southwest train.

Figure 3 Taxonomic profiling of RBC microbial community by the hidden Markov model (HMM) for the RNA polymerase beta subunit (*rpoB*). Stacked bars represent the relative abundance of unassembled *rpoB* metagenomic reads classified at the genus level. Genera at greater than or equal to 1% relative abundance are shown. Dark orange and green shaded bars represent nitrifiers. NE, northeast train; NW, northwest train; SE, southeast train; SW, southwest train.

Figure 4 Functional gene profiling of nitrifiers in RBC microbial communities by hidden Markov models (HMMs). Stacked bars represent the relative abundance of functional genes from unassembled metagenomic reads, classified to the genus level (B, C) or phylum level (A). Within each sample, length-normalized HMM hits for each functional gene were divided by the total number of length-normalized HMM hits for *rpoB.* Blue, green, and orange bars represent nitrifiers. Other taxa were classified to the genus level, but collapsed to the phylum level for clarity, and are shown in grey. Taxa shown are at greater than or equal to 1% relative abundance at the genus level relative to the HMM used. Taxa at less than 1% abundance are represented by the uppermost grey bar. (A) *amoA* gene of ammonia-oxidizing archaea. No HMM hits were found for sample NE1 June 2010. (B) *amoA* gene of ammonia-oxidizing bacteria and comammox bacteria. (C) *nxrB* gene of *Nitrospira*. NE, northeast train; NW, northwest train; SE, southeast train; SW, southwest train; *amoA*, ammonia monooxygenase subunit A gene; *nxrB*, nitrite oxidoreductase

Figure 5 Phylogeny and abundance distribution of *Nitrospira amoA* sequences. Phylogenetic analysis of *Nitrospira* AmoA from assembled contigs of all assembly combinations was inferred using maximum likelihood analysis. Bootstrap values over 60% are indicated with a grey dot, and those at 100% are shown with a black dot. The tree is drawn to scale, with branch lengths measured in the number of amino acid substitutions per site. AmoA sequence groups from RBCs are shown in blue. Cultivated comammox species are shown in green. Comammox AmoA sequences from environmental surveys are shown in black. *Methylocystis* sp. PmoA was used as an outgroup. AmoA cluster names followed by an asterisk indicate that the AmoA-encoding gene was assigned to a genome bin with ≥75% completeness and <25% contamination. The

average coverage of the bins within each sample, based on read mapping (see Methods), is depicted in the heatmap.

Table 1 Summary of dereplicated high quality *Nitrospira* bins with ≥75% completion and <25% contamination. Bins included were "winning" genomes for their cluster, as determined by a dRep analysis.

	SW1.007	CA-NE8-2010.42	Upper-2016.254	JNE1.001	CA-NE1-2010.95	CA-NE8-2010.148	Upper-2016.235	NW1.001	SNE8.002	SE8.002	SW.74	NE8.001	CA-NE8-2010.156	CA-NE8-2010.193	CA-NE8-2010.35	SW1.001
Completeness (%)	94.0	92.2	86.8	78.0	87.3	79.7	75.6	89.9	76.8	91.4	78.1	92.2	96.8	87.0	95.9	87.8
Contamination (%)	1.9	0.9	1.1	4.6	1.6	1.3	4.4	2.8	5.8	21.6	2.3	2.7	2.7	5.8	3.8	4.0
dRep cluster	1_1	2_1	3_1	4_0	5_0	6_1	7_0	8_1	8_2	8_3	9_1	10_1	11_1	12_1	13_1	14_1
Genome size (Mbp)	4.1	4.0	3.0	2.9	3.7	2.4	4.7	3.0	2.5	5.2	2.6	4.0	4.4	4.1	3.9	3.4
Number of contigs	310	57	471	628	526	561	531	397	466	634	505	317	159	341	106	429
Bin N50	24075	125286	8041	8005	9186	5233	17866	13033	8324	23758	5942	26973	48920	18888	117006	12389
GC content (%)	49.8	59.5	58.8	58.9	58.6	55.3	59.2	55.1	54.5	57.9	54.4	54.9	54.4	55.1	54.9	55.2
amoA	-	-	-	-	-	-	-	+	+	-	-	+	-	+	-	-
amoB	-	-	-	-	-	-	-	+	+	-	-	+	+	+	+	+
amoC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
amt	-	-	-	-	-	-	-	+	-	-	+	+	+	+	+	+
hao	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+
nxrB	-	-	-	-	+	-	-	-	+	+	-	+	-	-	+	-
ure	-	-	+	-	-	+/-	+	+	+	+/-	+	+	+	+	+	+
ccm	-	-	-	-	-	+	-	+	+	+	+	+	+	+	+	+
cyc	-	-	-	-	-	-	-	+	-	-	-	+	+	+	+	+
rpoB	+	+	+	+	-	+	-	-	+	+	+	+	+	-	+	+
hyb	+/-	-	+/-	-	-	-	+/-	+/-	+/-	-	-	+/-	+/-	+/-	-	-
hyd	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
hyp	+/-	-	+/-	-	-	+/-	+/-	+/-	+/-	+/-	-	+/-	+/-	+/-	-	-
fdh	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-	-
cynS	+	+	+	+	+	-	+	-	-	+	-	-	-	-	+	-

+/- indicates only some of the subunits are present

amo, ammonia monooxygenase; *amt*, ammonia transporter; *hao,* hydroxylamine oxidoreductase; *nxrB*, nitrite oxidoreductase; *ure*, urease; *ccm*, cytochrome c biogenesis protein; *cyc*, cytochrome c; *rpoB, RNA polymerase; hyb* and *hyd*, group 3 [Ni-Fe] sulfur-reducing hydrogenase; *hyp*, hydrogenase accessory protein; *fdh*, formate dehydrogenase; *cyn*, cyanate hydratase

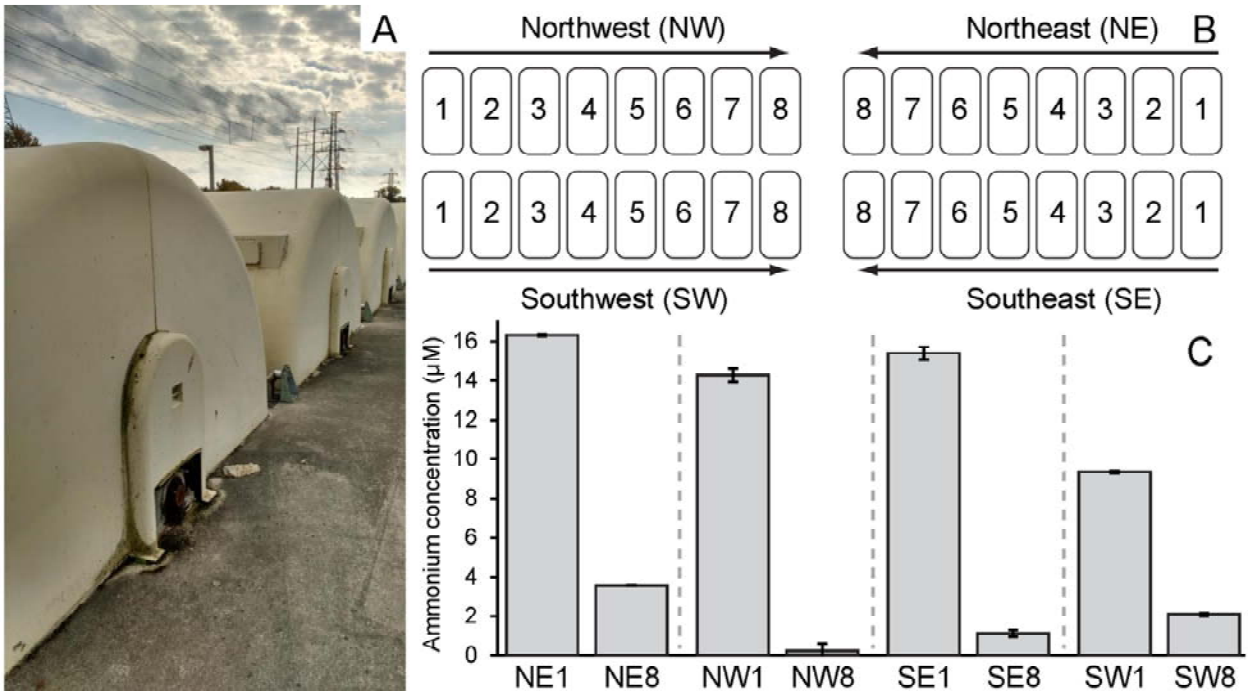
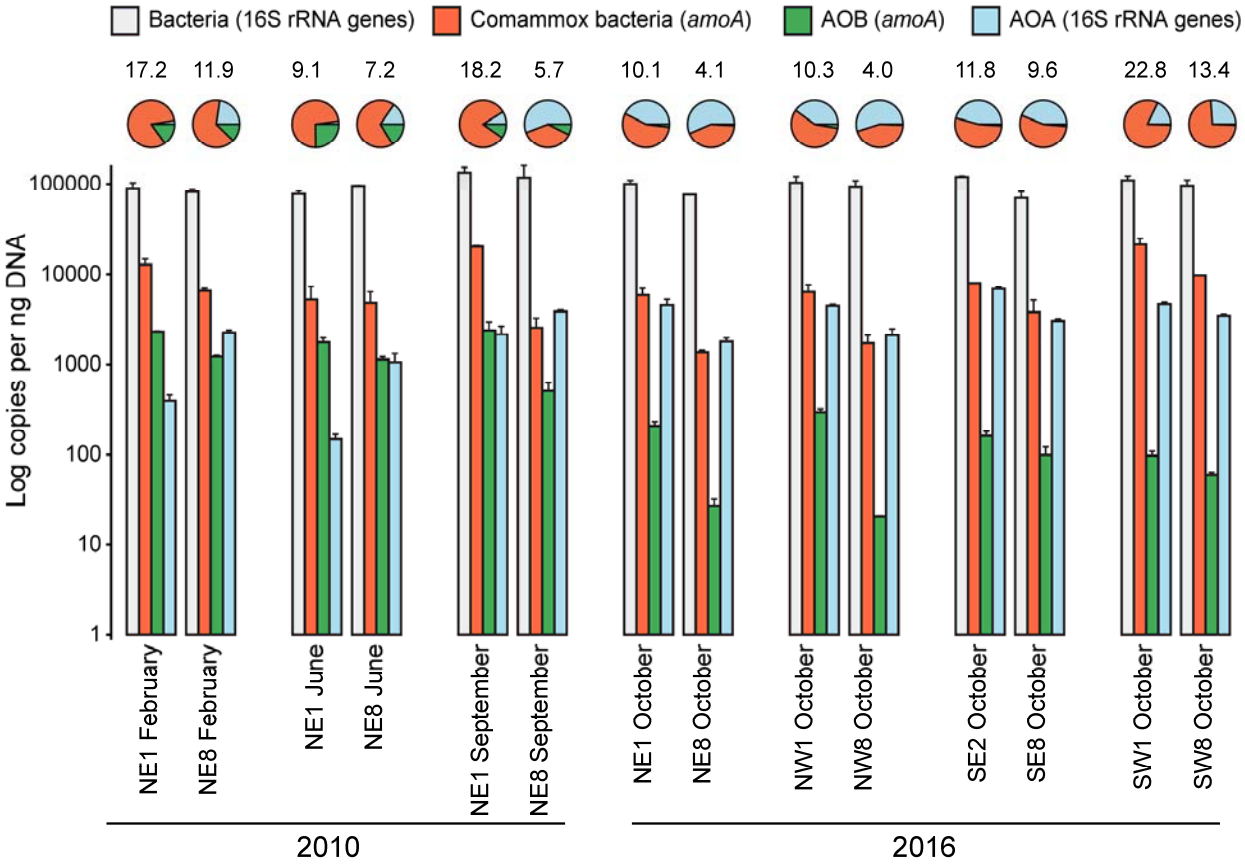

A
Northwest (NW)
Northeast (NE)
B
1 2 3 4 5 6 7 8
8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8
8 7 6 5 4 3 2 1
Southwest (SW)
Southeast (SE)
C
Ammonium concentration (μM)
0
2
4
6
8
10
12
14
16
NE1
NE8
NW1
NW8
SE1
SE8
SW1
SW8

Figure 1


Bacteria (16S rRNA genes)
Comammox bacteria (amoA)
AOB (amoA)
AOA (16S rRNA genes)
17.2
11.9
9.1
7.2
18.2
5.7
10.1
4.1
10.3
4.0
11.8
9.6
22.8
13.4
Log copies per ng DNA
100000
10000
1000
100
10
1
NE1 February
NE8 February
NE1 June
NE8 June
NE1 September
NE8 September
NE1 October
NE8 October
NW1 October
NW8 October
SE2 October
SE8 October
SW1 October
SW8 October
2010
2016

735 **Figure 2**

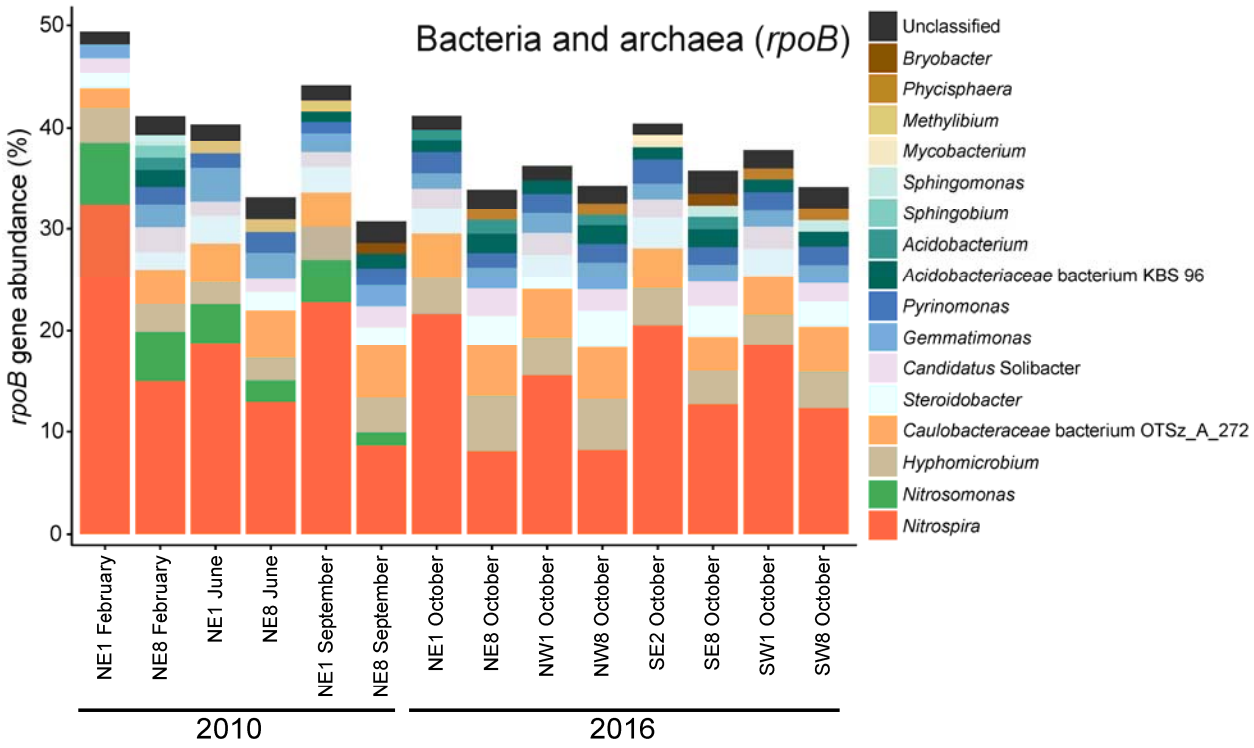

Bacteria and archaea (rpoB)
rpoB gene abundance (%)
50
40
30
20
10
0
NE1 February
NE8 February
NE1 June
NE8 June
NE1 September
NE8 September
NE1 October
NE8 October
NW1 October
NW8 October
SE2 October
SE8 October
SW1 October
SW8 October
2010
2016
Unclassified
Bryobacter
Phycisphaera
Methylibium
Mycobacterium
Sphingomonas
Sphingobium
Acidobacterium
Acidobacteriaceae bacterium KBS 96
Pyrinomonas
Gemmatimonas
Candidatus Solibacter
Steroidobacter
Caulobacteraceae bacterium OTSz_A_272
Hyphomicrobium
Nitrosomonas
Nitrospira

Figure 3

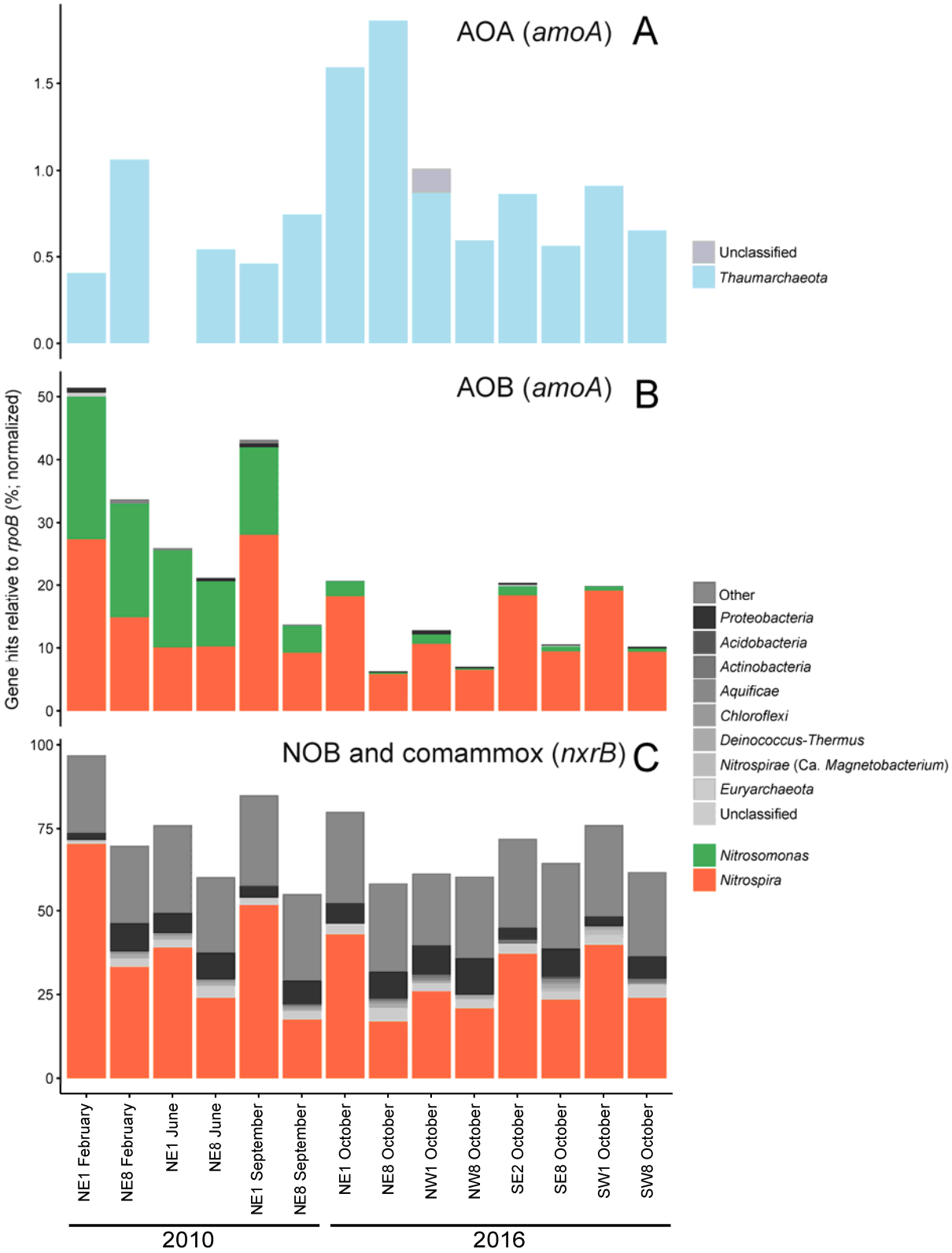

AOA (amoA) A
1.5
1.0
0.5
0.0
Unclassified
Thaumarchaeota
AOB (amoA) B
Gene hits relative to rpoB (%; normalized)
50
40
30
20
10
0
Other
Proteobacteria
Acidobacteria
Actinobacteria
Aquificae
Chloroflexi
Deinococcus-Thermus
Nitrospirae (Ca. Magnetobacterium)
Euryarchaeota
Unclassified
Nitrosomonas
Nitrospira
NOB and comammox (nxrB) C
100
75
50
25
0
NE1 February
NE8 February
NE1 June
NE8 June
NE1 September
NE8 September
NE1 October
NE8 October
NW1 October
NW8 October
SE2 October
SE8 October
SW1 October
SW8 October
2010
2016

Figure 4

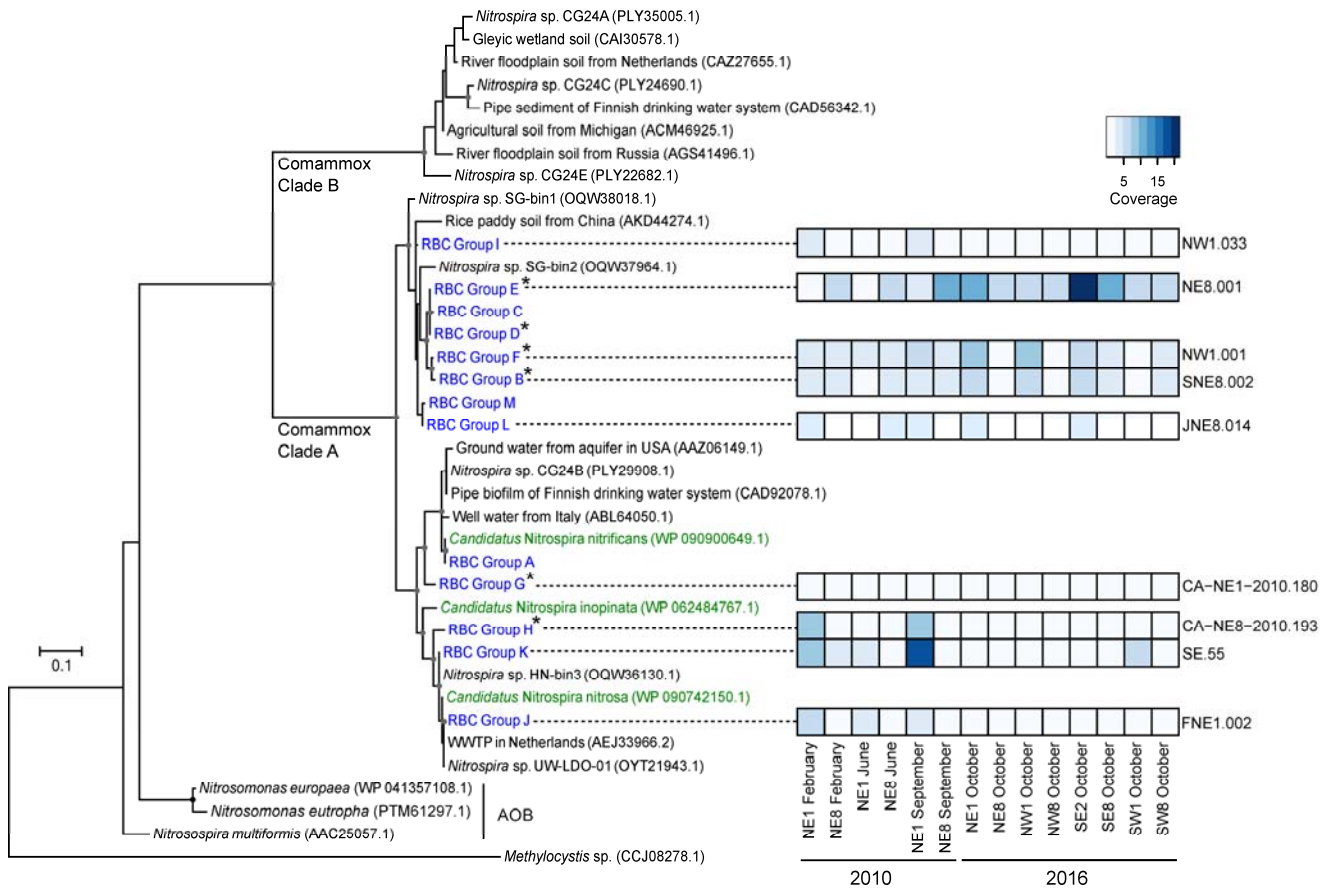

Nitrospira sp. CG24A (PLY35005.1)
Gleyic wetland soil (CAI30578.1)
River floodplain soil from Netherlands (CAZ27655.1)
Nitrospira sp. CG24C (PLY24690.1)
Pipe sediment of Finnish drinking water system (CAD56342.1)
Agricultural soil from Michigan (ACM46925.1)
River floodplain soil from Russia (AGS41496.1)
Nitrospira sp. CG24E (PLY22682.1)
Comammox Clade B
Nitrospira sp. SG-bin1 (OQW38018.1)
Rice paddy soil from China (AKD44274.1)
RBC Group I
Nitrospira sp. SG-bin2 (OQW37964.1)
RBC Group E*
RBC Group C
RBC Group D*
RBC Group F*
RBC Group B*
RBC Group M
RBC Group L
Comammox Clade A
Ground water from aquifer in USA (AAZ06149.1)
Nitrospira sp. CG24B (PLY29908.1)
Pipe biofilm of Finnish drinking water system (CAD92078.1)
Well water from Italy (ABL64050.1)
Candidatus Nitrospira nitrificans (WP 090900649.1)
RBC Group A
RBC Group G*
Candidatus Nitrospira inopinata (WP 062484767.1)
RBC Group H*
RBC Group K
Nitrospira sp. HN-bin3 (OQW36130.1)
Candidatus Nitrospira nitrosa (WP 090742150.1)
RBC Group J
WWTP in Netherlands (AEJ33966.2)
Nitrospira sp. UW-LDO-01 (OYT21943.1)
Nitrosomonas europaea (WP 041357108.1)
Nitrosomonas eutropha (PTM61297.1)
AOB
Nitrosospira multiformis (AAC25057.1)
Methylocystis sp. (CCJ08278.1)
0.1
5 15
Coverage
NW1.033
NE8.001
NW1.001
SNE8.002
JNE8.014
CA-NE1-2010.180
CA-NE8-2010.193
SE.55
FNE1.002
NE1 February
NE8 February
NE1 June
NE8 June
NE1 September
NE8 September
NE1 October
NE8 October
NW1 October
NW8 October
SE2 October
SE8 October
SW1 October
SW8 October
2010
2016

Figure 5