Design and Assessment of Species-Level qPCR Primers Targeting Comammox

Natalie K. Beach and Daniel R. Noguera*

Department of Civil and Environmental Engineering, University of Wisconsin–Madison, Madison, WI, United States

Published PCR primers targeting the ammonia monooxygenase gene (*amoA*) were applied to samples from activated sludge systems operated with low dissolved oxygen (DO) to quantify total and clade-level *Nitrospira* that perform complete ammonium oxidation (comammox); however, we found these existing primers resulted in significant artifact-associated non-target amplification. This not only overestimated comammox *amoA* copies but also resulted in numerous false positive detections in the environmental samples tested, as confirmed by gel electrophoresis. Therefore, instead of attempting to quantify comammox diversity, we focused on accurately quantifying the candidate comammox species. We designed specific and sensitive primers targeting 3 candidate species: *Candidatus* (*Ca.*) Nitrospira nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans. The primers were tested with *amoA* templates of these candidate species and used to quantify comammox at the species level in low DO activated sludge systems. We found that comammox related to *Ca.* N. nitrosa were present and abundant in the majority of samples from low DO bioreactors and were not detected in samples from a high DO system. In addition, the greatest abundance of *Ca.* N. nitrosa was found in bioreactors operated with a long solids retention time. *Ca.* N. inopinata and *Ca.* N. nitrificans were only detected sporadically in these samples, indicating a minor role of these comammox in nitrification under low DO conditions.

Keywords: comammox, *Nitrospira*, nitrification, low dissolved oxygen, biological nutrient removal, qPCR, real-time PCR, PCR primers

INTRODUCTION

The oxidation of ammonium via nitrite to nitrate (i.e., nitrification) was historically considered a two-step process completed by phylogenetically distinct ammonia oxidizers and nitrite oxidizers. Although the complete oxidation of ammonium to nitrate by a single organism was theoretically possible, no microorganism with the ability to carry out both steps had been identified until late 2015 (Daims et al., 2015; van Kessel et al., 2015). These complete ammonium oxidizing (comammox) bacteria belong to the genus *Nitrospira*, which was known to contain only nitrite-oxidizing bacteria (NOB); therefore, an entire group of microorganisms with the ability to oxidize ammonia had been disguised for years (Daims et al., 2015; van Kessel et al., 2015). The discovery of comammox bacteria has redefined a key component of the global nitrogen cycle, initiating a renewed focus of recent nitrogen cycling research in environmental biotechnology (Chao et al., 2016; Palomo et al., 2016; Pinto et al., 2016; Stadler and Love, 2016; Bartelme et al., 2017; Fowler et al., 2017; Hu and He, 2017).

Within the wastewater treatment industry, the application of nitrification under very low dissolved oxygen (DO) concentrations (below 0.2 mg O_2/L) is an exciting energy-saving approach to biological nutrient removal (BNR) (Park and Noguera, 2004; Paredes et al., 2007; Fitzgerald et al., 2015). Previous attempts to identify the key microorganisms responsible for nitrification in low DO bioreactors have been inconclusive, with the presence of known ammonia oxidizing bacteria (AOB) and ammonia oxidizing archaea (AOA) unable to explain observed nitrification rates (Fitzgerald et al., 2015). Thus, the recent discovery of comammox in a variety of environments (Daims et al., 2015; van Kessel et al., 2015) prompted us to investigate their presence in low DO BNR systems.

The gene encoding the alpha-subunit of ammonia monooxygenase (*amoA*) has been one of the most widely used markers for detection and quantification of AOB and AOA, as it facilitates functional analysis and reconstruction of phylogenetic relationships (Rotthauwe et al., 1997; Purkhold et al., 2000; Koops et al., 2006; Pester et al., 2012). The *amoA* of comammox is distinguishable from *amoA* sequences of AOB and AOA, and thus, it can be used to detect the presence of comammox in environmental samples (Daims et al., 2015; van Kessel et al., 2015; Pinto et al., 2016; Camejo et al., 2017). Recent descriptions of PCR primers targeting the *amoA* gene of comammox include a primer pair designed specifically for *Candidatus* (*Ca.*) Nitrospira inopinata (Daims et al., 2015), a primer pair designed specifically for a comammox-like clone within a freshwater aquaculture system (Bartelme et al., 2017), a primer collection that differentiates two broad clades of comammox within the *Nitrospira* genus (Pjevac et al., 2017), and a highly degenerate primer pair attempting to encompass all comammox within *Nitrospira* (Fowler et al., 2017). To date, three candidate comammox species have been described in the literature, namely *Ca.* N. nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans (Daims et al., 2015; van Kessel et al., 2015). However, quantifying the contribution of these candidate species to the comammox community in environmental samples has not been possible because of the lack of specific qPCR primer sets. To overcome this limitation, we designed a set of highly specific non-degenerate primers for the independent detection of *Ca.* N. nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans. We used these primers to evaluate the contribution of these species to the comammox population in samples from BNR plants operated at low DO conditions.

MATERIALS AND METHODS

Sample Collection, Processing, and DNA Extraction

Environmental samples used in this study originated from four low DO nitrifying bioreactors: a laboratory-scale sequencing batch reactor (L_SBR), a pilot-scale sequencing batch reactor (P_SBR), a pilot scale continuous flow (P_CF) reactor simulating a University of Cape Town configuration without nitrate recycle, and a full-scale wastewater treatment plant (WWTP) from the Trinity River Authority (TRA) Central Region (Arlington,

TX, United States) (**Table 1**). For comparison, samples were also collected from the full-scale Nine Springs WWTP (NS) at the Madison Metropolitan Sewerage District (Madison, WI, United States), which operates with typical high DO conditions (**Table 1**). Grab samples with varying volumes were collected from every reactor and were either centrifuged to form a biomass pellet, with supernatant discarded, and the pellet frozen at −80°C (P_SBR, P_CF, TRA, and NS) or were saved in a glycerol mixture prior to placement in a −80°C freezer (L_SBR). Biomass samples were stored at −80°C until DNA extraction. All DNA was extracted using DNeasy® PowerSoil® DNA Isolation Kit (Qiagen, Hilden, Germany) following the manufacturer's directions. DNA was quantified with a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, United States) and the purity ratio, or ratio of absorbance at 260 and 280 nm, was determined with a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, United States). DNA samples were stored at −20°C until further processing.

Design of Primers to Detect Comammox Ammonia Monooxygenase Gene *amoA*

A collection of full-length *amoA* and full-length particulate methane monooxygenase (*pmoA*) gene sequences, obtained from the National Center for Biotechnology Information (NCBI) GenBank database (NCBI Resource Coordinators, 2017), were used for primer design. The full-length sequences included in the design are indicated by bold text in the phylogenetic tree in **Figure 3**, with the corresponding accession number in parenthesis. All *amoA* and *pmoA* sequences were aligned using the 'AlignSeqs' command in the DECIPHER "R" package (Wright, 2015, 2016). This aligned database was then submitted to DECIPHER's Design Primers web tool (Wright et al., 2014). Sequences corresponding to *Ca.* N. nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans were individually selected as target groups for primer design, which used the following parameters: primer length ranging from 17 – 26 nucleotides with up to 1 permutation, PCR product amplicon length of 100 – 450 bp, 100% target group coverage, and without the Taq 3′-end Model option. The primer design tool also used the same reaction conditions in all cases: [Na^+] 70 mM, [Mg^{2+}] 3 mM, [dNTPs] 0.8 mM, annealing temperature (T_a) of 64°C, [primers] 400 nM. Amplification products were verified by agarose (2%) gel electrophoresis with GelRed™ Nucleic Acid Gel Stain (Biotium, Freemont, CA, United States).

Quantitative Real-Time Polymerase Chain Reaction (qPCR)

Quantification of total 16S rRNA genes, total comammox *amoA* genes, clades A and B comammox *amoA*, as well as *Ca.* N. nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans *amoA* in each DNA sample was carried out by qPCR. All qPCR assays were performed on a Roche LightCycler® 480 high-throughput real-time PCR system using white LightCycler® 480 multiwell plates and the associated LightCycler® 480 sealing foils (Roche Molecular Systems, Inc., Pleasanton, CA, United States). All environmental DNA samples were diluted to 10 ng/μL. Although not specifically

TABLE 1 | Bioreactor sample characteristics.

Bioreactor (Sample ID)	Location	Sample Dates[a]	Configuration at time of sampling	Size (gpd)	SRT (days)	Reference
Lab-scale sequencing batch reactor (L_SBR)	Madison, WI, United States	3/7/2013 (1), 3/18/2013 (2), 4/17/2013 (3), 10/3/2014 (4)	Sequencing batch reactor (SBR) operated with dissolved oxygen (DO) < 0.2 mg O_2/L	0.5	80	Camejo et al., 2016
Pilot-scale sequencing batch reactor (P_SBR)	Madison, WI, United States	9/10/2015 (1), 11/10/2015 (2), 12/15/2015 (3), 4/25/2016 (4)	Sequencing batch reactor (SBR) operated with DO < 0.7 mg O_2/L	130	80	Camejo et al., 2016
Continuous flow pilot plant (P_CF)	Madison, WI, United States	9/9/2015 (1), 11/18/2015 (2), 12/30/2015 (3), 4/26/2016 (4)	Modified University of Capetown (UCT) operated with DO < 0.5 mg O_2/L	1200	10	Keene et al., 2017
Trinity River Authority Central Region Wastewater System Treatment Plant (TRA)[b]	Arlington, TX, United States	8/9/2016 (1), 7/28/2017[b] (2), 7/28/2017 (3)	Anaerobic Aerobic (AO) operated with three zones. Zone 1: DO < 0.6 mg O_2/L Zone 2: DO < 0.9 mg O_2/L Zone 3: DO < 1.4 mg O_2/L	123 M	10	This study
Madison Metropolitan Sewerage District Nine Springs Plant (NS)	Madison, WI, United States	9/9/2015 (1), 11/18/2015 (2), 12/30/2015 (3), 4/26/2016 (4)	Modified University of Capetown (UCT) operated with DO > 2.0 mg O_2/L	40 M	10	Keene et al., 2017

[a]Number in parentheses after each date is incorporated after the sample ID.
[b]TRA sample 2 and 3 were collected from different aeration basins.

studied here, the applied sample dilutions helped reduce any potential problems with PCR inhibition. All qPCR assays were prepared in Bio-Rad 2x iQ™ SYBR® Green Supermix (Bio-Rad, Hercules, CA, United States), containing 50 U/mL iTaq DNA polymerase, 1.6 mM dNTPs, 100 mM KCl, 40 mM Tris-HCl, 6 mM $MgCl_2$, 20 nM fluorescein, and stabilizers (10 µL per reaction). Triplicate reactions were prepared for each sample.

Amplification of total comammox *amoA* was performed according to Fowler et al. (2017), with each reaction containing Bio-Rad 2x iQ™ SYBR® Green Supermix (10 µL), nuclease free water (4.4 µL), environmental DNA or standard (4 µL), and PCR primers (0.8 µL of 12.5 µM each primer) added for a final volume of 20 µL per reaction. The thermal cycling protocol for the total comammox primers was as follows: initial denaturation step at 94°C for 5 min, followed by 40 cycles of initial denaturation at 94°C for 30 s, annealing at 48°C for 30 s, and extension at 72°C for 1 min.

Amplification of clade A and clade B comammox *amoA* was performed using the equimolar primer mixtures according to Pjevac et al. (2017) (**Table 2**). Four equimolar primer mixtures were created for clade A and clade B forward and reverse primers, by combining 62.5 µL of each of the six individual primer stocks (100 mM) together with 125 µL nuclease free water to a final working volume of 500 µL and final concentration of 12.5 µM each primer. Finally, reactions were prepared in Bio-Rad 2x iQ™ SYBR® Green Supermix (10 µL), nuclease free water (4.4 µL), environmental DNA or standard (4 µL), and equimolar PCR primer mixture (0.8 µL of 12.5 µM each primer) were added to each reaction for a final volume of 20 µL per reaction. The thermal cycling protocol for the clade-level primers was as

follows: initial denaturation step at 95°C for 3 min, followed by 45 cycles of initial denaturation at 95°C for 30 s, annealing at 52°C for 45 s, and extension at 72°C for 1 min.

For the novel *amoA* primers designed in this study (**Table 2**), nuclease free water (4.4 µL), environmental DNA or standard (4 µL), and PCR primers (0.8 µL of 10 µM each primer) were added to each reaction for a final volume of 20 µL per reaction. The thermal cycling protocol for qPCR using the novel *amoA* primers designed in this study (**Table 2**) was as follows: initial denaturation step at 95°C for 10 min, followed by 45 cycles of initial denaturation at 95°C for 10 s, annealing at 64°C for 30 s, and extension at 72°C for 30 s. Fluorescence was measured at 72°C for amplicon quantification. After amplification, an amplicon melting curve was recorded in 0.25°C steps between 65 and 97°C. Melting peaks were obtained by plotting the negative first derivative of fluorescence against temperature. Although 30 cycles is typically sufficient for quantification of targets in qPCR, the thermal cycling was extended to evaluate potential non-specific amplification with the newly designed primers (Wright et al., 2014). Finally, amplification of total 16S rRNA was performed using the 16S rRNA-targeted primer pair 341f/785r according to Thijs et al. (2017) (**Table 2**).

Standard Preparation for qPCR

Full-length *amoA* from *Candidatus* Nitrospira nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans were used as standards for the new species-specific comammox assay as well as the total comammox and clade A comammox assays. Full-length *amoA* from *Nitrospira sp. CG24_E* (Palomo et al., 2018) was used as the standard for the clade B comammox assay. These individual *amoA* standards were generated from synthetic gene

TABLE 2 | Primers used for qPCR.

Target gene	Primer name	Forward primer (5' - 3')	Reverse primer (5' - 3')	T_a(°C)	Reference
Ca. Nitrospira nitrosa *amoA*	Nitrosa amoA-469F/812R	GCGATTCTGTTTTATCCCAGCAAC	CCGTGTGCTAACGTGGCG		
Ca. Nitrospira inopinata *amoA*	Inopinata amoA-410F/815R	TCACCTTGTTGCTAACTAGAAACTGG	TCCGCGTGAGCCAATGT	64	This study
Ca. Nitrospira nitrificans *amoA*	Nitrificans amoA-463F/836R	ATGTTCGCGGCACTGTT	CCAGAAAGTTTAGCTTTGTCGCCT		
Comammox *Nitrospira* clade A *amoA*	comaA-244f_a/659r_a	TACAACTGGGTGAACTA	AGATCATGGTGCTATG	52	Pjevac et al., 2017
	comaA-244f_b/659r_b	TATAACTGGGTGAACTA	AAATCATGGTGCTATG		
	comaA-244f_c/659r_c	TACAATTGGGTGAACTA	AGATCATGGTGCTGTG		
	comaA-244f_d/659r_d	TACAACTGGGTCAACTA	AAATCATGGTGCTGTG		
	comaA-244f_e/659r_e	TACAACTGGGTCAATTA	AGATCATCGTGCTGTG		
	comaA-244f_f/659r_f	TATAACTGGGTCAATTA	AAATCATCGTGCTGTG		
Comammox *Nitrospira* clade B *amoA*	comaB-244f_a/659r_a	TAYTTCTGGACGTTCTA	ARATCCAGACGGTGTG	52	Pjevac et al., 2017
	comaB-244f_b/659r_b	TAYTTCTGGACATTCTA	ARATCCAAACGGTGTG		
	comaB-244f_c/659r_c	TACTTCTGGACTTTCTA	ARATCCAGACAGTGTG		
	comaB-244f_d/659r_d	TAYTTCTGGACGTTTTA	ARATCCAAACAGTGTG		
	comaB-244f_e/659r_e	TAYTTCTGGACATTTTA	AGATCCAGACTGTGTG		
	comaB-244f_f/659r_f	TACTTCTGGACCTTCTA	AGATCCAAACAGTGTG		
Total comammox *Nitrospira amoA*	Ntsp-amoA 162F/359R	GGATTTCTGGNTSGATTGGA	WAGTTNGACCACCASTACCA	48	Fowler et al., 2017
Total *16S rRNA*	16S-341f/785r	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	53	Klindworth et al., 2013; Thijs et al., 2017

plasmid cloning vectors (Integrated DNA Technologies, Inc., Coralville, IA, United States) transformed into One Shot[TM] TOP10 Competent *Escherichia coli* (Life Technologies, Carlsbad, CA, United States) with the TOPO-TA cloning kit (Invitrogen, Karlsruhe, Germany) which uses the pCR[TM]4-TOPO® TA vector. The cloned plasmids were subjected to amplification with M13 primers and product sizes verified with gel electrophoresis. The M13-PCR products were purified using the Qiagen PCR Purification Kit (Qiagen, Hilden, Germany) and quantified using a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, United States). The purity ratio of the standard was determined with a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, United States).

Standard gene copy number was calculated from the DNA concentration. The standards were diluted in a 10-fold series over seven orders of magnitude (approximately 10^1 to 10^7 copies) and were included in triplicate with each set of samples. A 95% confidence interval and standard deviation was calculated for the fractional PCR cycle used for quantification (i.e., quantification cycle or C_q) for each set of standard replicates. A regression line fit to the C_q and the known concentration of each standard was used to generate a standard calibration curve for each assay. Next, the standard replicate statistics and assay-specific standard calibration curve were used to approximate a limit of detection (LOD) and limit of quantification (LOQ) according to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE Guidelines) described by Bustin et al. (2009) and equations adapted from Armbruster and Pry (2008). Median LOD, median LOQ, and corresponding 95% confidence intervals are reported for each assay (**Table 3**). External standard curve amplification efficiencies were calculated

as $E = (10^{\frac{-1}{m}} - 1) \times 100$, where m represents the slope of the standard curve.

Gradient PCR

Gradient PCR was performed with the total comammox (Fowler et al., 2017) and clade-level comammox (Pjevac et al., 2017) *amoA* primers with environmental DNA samples diluted to three concentrations to determine if an optimal annealing temperature and DNA concentration exists to eliminate or at least minimize the observed non-specific amplification. Gradient PCR was performed on a Mastercycler® nexus gradient PCR cycler (Eppendorf North America, Hauppauge, NY, United States) with Eppendorf twin.tec® 96-well LoBind PCR plates and Eppendorf PCR sealing foils. Reaction mixtures and thermal cycling protocols were identical to the respective qPCR assay, with the exception that annealing temperatures were set to vary across each PCR plate. For the total comammox primers, annealing temperatures were set in small increments between 47.8 and 56°C. For the clade-level comammox primers, annealing temperatures were set in small increments between 51.8 and 60°C. Two samples were selected for each assay and were selected based on the presence of both the target and non-specific amplifications after qPCR. The total comammox gradient assay included DNA from P_SBR-4 and NS-4. The clade A and clade B comammox gradient assays included DNA from P_CF-4 and NS-4. All environmental DNA samples were diluted to 40 ng DNA/reaction, 20 ng DNA/reaction, and 1 ng DNA/reaction. Amplification products were then visualized with agarose (2%) gel electrophoresis with GelRed[TM] Nucleic Acid Gel Stain (Biotium, Freemont, CA, United States).

TABLE 3 | qPCR performance with three distinct comammox *amoA* standard templates.

	Primer set	Amplicon size (bp)	Amplicon GC content (%)	Average T_m (°C)	Log_{10} linear range, log(SQ)*	Average PCR efficiency (%)	Median LOD** (copies/ng DNA)	Median LOQ** (copies/ng DNA)
Comammox *Nitrospira* clade A	**Standard template: *N. nitrosa amoA***							
	Nitrosa amoA-469F/812R[a]	344	57	88.8 ± 0.08	1 - 7	96 ± 3.0	M = 11, 95% CI [9.3, 13]	M = 13, 95% CI [10, 17]
	Ntsp-amoA 162F/369R[b]	198	54	86.1 ± 0.08	3 - 7	91 ± 2.6	M = 20, 95% CI [8.6, 48]	M = 28, 95% CI [11, 72]
	comaA-244f/659r[c]	415	58	88.9 ± 0.09	2 - 7	95 ± 2.7	M = 38, 95% CI [13, 110]	M = 40, 95% CI [14, 120]
	Standard template: *N. inopinata amoA*							
	Inopinata amoA-410F/815R[a]	406	49	85.6 ± 0.05	1 - 7	90 ± 3.0	M = 7.0, 95% CI [4.0, 12]	M = 14, 95% CI [7.0, 29]
	Ntsp-amoA 162F/369R[b]	198	49	84.4 ± 0.60	3 - 7	82 ± 3.4	M = 20, 95% CI [8.6, 48]	M = 28, 95% CI [11, 72]
	comaA-244f/659r[c]	415	50	86.0 ± 0.10	3 - 7	86 ± 2.5	M = 38, 95% CI [13, 110]	M = 40, 95% CI [14, 120]
	Standard template: *N. nitrificans amoA*							
	Nitrificans amoA-463F/836R[a]	374	52	86.7 ± 0.10	1 - 7	92 ± 2.4	M = 1.3, 95% CI [0.9, 2.0]	M = 3.8, 95% CI [1.3, 11]
	Ntsp-amoA 162F/369R[b]	198	57	87.3 ± 0.27	2 - 7	96 ± 1.8	M = 20, 95% CI [8.6, 48]	M = 28, 95% CI [11, 72]
	comaA-244f/659r[c]	415	56	88.2 ± 0.07	2 - 7	93 ± 2.9	M = 38, 95% CI [13, 110]	M = 40, 95% CI [14, 120]

*Average parameters are presented with standard deviation. PCR efficiency is reported for the linear range of the standard calibration curve. *SQ, starting quantity in copies per reaction. **M, median; CI, confidence interval [upper limit, lower limit]; [a] This study; [b] Fowler et al., 2017. [c] Pjevac et al., 2017.*

Phylogenetic Tree Construction

The database of aligned *amoA* and *pmoA* sequences was expanded by including additional comammox *amoA* sequences obtained from metagenomes available on the JGI Integrated Microbial Genomes (IMG) database (Markowitz et al., 2012) and available comammox *Nitrospira* bins extracted with mmgenome[1] (Karst et al., 2016). The expanded database was utilized for *in silico* primer analyses and to construct a consensus phylogenetic tree in Geneious v. 9.1.2 (Kearse et al., 2012) using the Neighbor-Joining method on a Tamura-Nei genetic distance model with bootstrap resampling and 20,000 trials.

Average Nucleotide Identity Calculation

To determine if comammox *Nitrospira* draft genomes belonged to the same species, we performed pairwise analyses of genome average nucleotide identity (ANI). Initially, open reading frames (ORFs) within each draft genome was predicted with Prodigal (Hyatt et al., 2010). Subsequently, the JSpecies Web Server (Richter et al., 2016) was used to calculate ANI based on BLAST+ (Camacho et al., 2009) from percent of aligned genome ORFs.
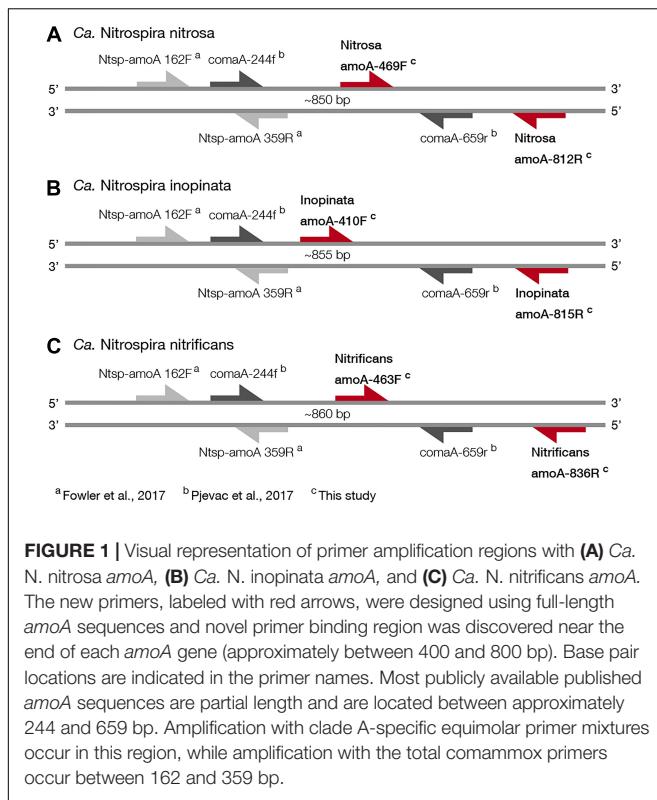
RESULTS

Total Comammox and Clade-Level Comammox Detection

For broad comammox *amoA* quantification, three published primer sets were compared using standards and nineteen environmental samples originating from five bioreactors (**Table 1**). The Ntsp-amoA 162F/359R primer set (Fowler et al., 2017), designed to target total comammox *Nitrospira amoA* (**Table 2**), amplifies a 198 bp fragment at primer binding regions between positions 162–182 and 339–359 bp (**Figure 1**). The comaA-244f/659R and comaB-244f/659r primer sets (Pjevac et al., 2017), designed to differentiate between clades A and B of comammox (**Table 2**), amplify a 415 bp fragment at *amoA* primer binding regions between positions 244–261 and 643–659 bp (**Figure 1**). The standard curves for these assays were linear over a minimum of five orders of magnitude and had high coefficients of determination ($R^2 > 0.990$); although linear range, amplification efficiency, LOD and LOQ varied depending on the comammox *amoA* standard used (**Table 3** and **Supplementary Table S1**). The results for the total comammox (Fowler et al., 2017) and clade-level comammox (Pjevac et al., 2017) assays are also summarized in decision matrices that were used to consistently evaluate qPCR results between assays (**Supplementary Tables S3–S5**).

In the total comammox assay, melting curves for the standards were narrow and unimodal above ∼10^1 – 10^2 copies (**Supplementary Figures S1A–C**). However, melting curves from the environmental samples (**Supplementary Figures S1D–H**) and the agarose gel of qPCR products (**Supplementary Figure S2**) showed non-specific amplifications resulting in three false positive detections (amplification did not correspond

---

[1]http://madsalbertsen.github.io/mmgenome/

FIGURE 1 | Visual representation of primer amplification regions with **(A)** *Ca.* N. nitrosa *amoA*, **(B)** *Ca.* N. inopinata *amoA,* and **(C)** *Ca.* N. nitrificans *amoA*. The new primers, labeled with red arrows, were designed using full-length *amoA* sequences and novel primer binding region was discovered near the end of each *amoA* gene (approximately between 400 and 800 bp). Base pair locations are indicated in the primer names. Most publicly available published *amoA* sequences are partial length and are located between approximately 244 and 659 bp. Amplification with clade A-specific equimolar primer mixtures occur in this region, while amplification with the total comammox primers occur between 162 and 359 bp.

to expected product size) and overestimation of comammox abundance (multiple products seen, including one with the expected product size) in six samples (**Figure 2A** and **Supplementary Table S3**).

In the clade A comammox assay, all standards with approximately 10^2 *amoA* copies or less presented two melting peaks, the expected target peak at approximately 88.9, 86.0, and 88.2°C for *Ca.* N. nitrosa*, Ca.* N. inopinata, and *Ca.* N. nitrificans *amoA*, respectively, plus a sizeable off-target peak at approximately 70°C (**Supplementary Figures S3A–C**). Higher copy numbers produced narrow, unimodal melting peaks at the expected T_m. The agarose gel performed with the qPCR products showed a single band for the standards; however, eight of the nineteen environmental samples contained multiple non-specific amplified products including a potential target amplicon (**Supplementary Figures S3D–H, S4**) leading to overestimated gene abundance (**Figure 2B** and **Supplementary Table S4**). Amplification of non-specific products in seven additional samples produced artifact-associated fluorescence that was well over the LOQ and translated into false positive predictions of the same order of magnitude as true positive detections (**Figure 2B**).

Clade B standards with approximately 10^3 *amoA* copies or less also contained two melting peaks, an off-target peak at approximately 72°C and the expected target peak at approximately 89°C (**Supplementary Figure S5A**). In the clade B assay, fifteen samples had a target melting peak at the expected T_m (**Supplementary Figures S5B–F**), however, the agarose gel revealed that these samples had abundant non-specific products that were approximately 300 bp long— smaller than the expected

415 bp target size (**Supplementary Figure S6**). Only one sample (P_CF-4) appeared to contain an amplicon of the expected size (**Supplementary Figure S6**), but also contained several other non-specific amplicons. Thus, none of the samples tested were confirmed to be a true positive for presence of clade B comammox (**Supplementary Table S5**).

The biggest challenge with the non-specific amplification observed with the total comammox and clade-level comammox primers was the significant artifact-associated fluorescence resulting in C_q values within the same range as samples with abundant comammox (see **Supplementary Tables S3–S5**), which did not allow us to eliminate samples based on the LOD or LOQ. Two factors that could have contributed to the non-specific amplifications observed with the broad-detection primers are the low design annealing temperatures (48°C for the total comammox primers and 52°C for the clade A and clade B comammox primers) and the total DNA concentration we used in each assay. Therefore, we performed gradient PCR experiments to evaluate whether adjustments to the primer annealing temperature or sample DNA concentration would improve the specificity of the total comammox primers (Fowler et al., 2017) and the clade A and clade B comammox primer sets (Pjevac et al., 2017). In our specific application, the gradient PCR experiments show that small adjustments to annealing temperature or sample dilutions did not make a meaningful difference in the non-specific amplifications. Non-specific products remained in the samples tested at higher annealing temperatures and at all three DNA concentrations tested (40 ng/reaction, 20 ng/reaction, and 1 ng/reaction) with all three published primer sets. In the gradient PCR with the total comammox primers, non-specific products are observed at approximately 300 and 600 bp in the P_SBR-4 sample and at approximately 150, 350, 600, and 700 bp in the NS-4 sample, and do not appear to be dependent on the total DNA concentrations tested in the reaction (**Supplementary Figure S7**). In the gradient PCR for the clade A comammox primers, non-specific products are observed at approximately 150 and 300 bp in the P_CF-4 sample and at approximately 250 and 300 bp in the NS-4 sample, and the amplification of these non-specific products do not appear to be dependent on the total DNA concentrations tested in the reaction (**Supplementary Figure S8**). In the gradient PCR for the clade B comammox primers (**Supplementary Figure S9**), multiple non-specific products were observed, with the most prominent non-specific product consistent with previous observations at approximately 300 bp.

Another challenge we faced when applying the broad-detection comammox primers to our samples was the inability to rely on melting curve analysis alone to differentiate between true positive detections, false positive detections, and overestimated detections. To successfully complete melting curve analysis post qPCR, the melting characteristics of an amplified standard (characterized by the both the shape of the melting peak and the temperature where the peak occurs or T_m) is compared to the melting peaks that appear in the environmental samples since the T_m considered a unique feature of a particular DNA fragment (Ririe et al., 1997; Wittwer and Kusakawa, 2004). However, this type of analysis can be difficult with primer sets that target a broad
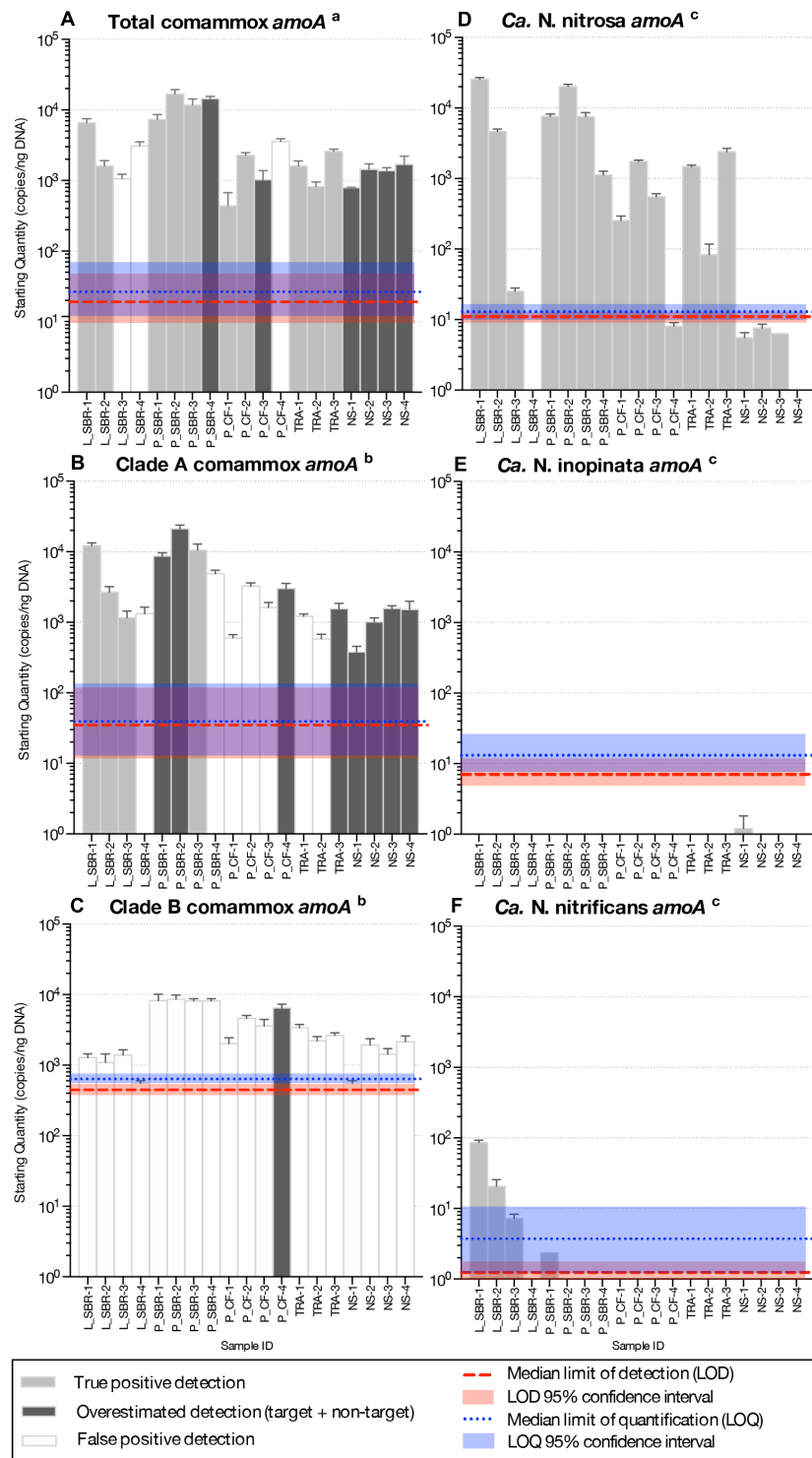
FIGURE 2 | Abundance of *amoA* genes (gene copies/ng DNA) on a log-scale from comammox *Nitrospira* using the **(A)** total comammox primers, **(B)** clade A primers **(C)** clade B primers, **(D)** the newly designed *Ca.* N. nitrosa specific primers **(E)** *Ca.* N. inopinata specific primers and **(F)** *Ca.* N. nitrificans specific primers from various time series samples originating from five bioreactors. LS, lab-scale sequencing batch reactor; P_SBR, pilot-scale sequencing batch reactor; P_CF, continuous flow pilot plant; TRA, Trinity River WWTP; NS, Nine Springs WWTP. Error bars show the standard deviation of the triplicate samples. The color of each column indicates whether there was a true positive detection, an overestimated detection, or a false positive detection. The median limit of detection and quantification for each assay are represented by horizontal red and blue dotted lines, respectively. The shaded regions surrounding these limits represent the upper and lower bounds of the 95% confidence interval.

number of sequences, especially when the primers are designed to amplify a region with highly variable GC content. In this study, we found that the total comammox primers (Fowler et al., 2017) amplify a region of *Ca.* N. nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans *amoA* with GC content at 54, 49, and 57%, respectively. This produced melting peaks for *Ca.* N. nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans *amoA* at a T_m of 86.1°C, 84.4°C, and 87.3°C, respectively. We also found that the clade A comammox primers (Pjevac et al., 2017) amplify a region of *Ca.* N. nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans *amoA* with GC content at 58, 50, and 56%, respectively. This, in turn, produced melting peaks for *Ca.* N. nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans *amoA* at a T_m of 88.9, 86.0, and 88.2°C, respectively. While it is possible to determine a T_m range based on *in silico* analysis of the known comammox diversity, this does not allow us to eliminate non-specific products from the analysis since some may contain similar melting characteristics as the target. For example, the melting peaks for P_CF-4 with the total comammox primers (**Supplementary Figure S1F**) fall within the T_m range set by the standards used in this study, but only contained a non-specific product between 400 and 500 bp long when verified with agarose gel electrophoresis (**Supplementary Figure S2**). Since the T_m is also influenced by the reaction conditions, one would need to include a much greater number of standards *in situ* to represent a greater variety of known comammox species in order to truly rely on melting curve analysis alone with these broad-detection assays. Since more species are continually discovered, relying on melting curve analysis alone to confirm positive detections will likely be difficult without an array of comammox *amoA* standards and will require confirmation with an agarose gel electrophoresis of amplified qPCR products.

Overall, the presence, frequency, and influence of unspecific amplification in the total comammox (Fowler et al., 2017) assay and in the clade A and clade B (Pjevac et al., 2017) comammox assays were significant and did not allow for accurate detection and quantification of comammox *Nitrospira* in the specific environmental samples tested in this study (**Figures 2A–C**), prompting us to design other options for specific comammox detection.

Design of Species-Specific Primer Sets

Since the broad range comammox primers did not provide satisfactory results with our environmental samples, we opted to focus our evaluations on the presence of specific comammox species. With this objective in mind, to detect and quantify comammox belonging to the *Candidatus Nitrospira* species currently described in the literature (*Ca.* N. nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans) we designed qPCR primer sets specifically targeting each of these species. For this design, we use a dataset of 27 full-length *amoA* gene sequences from AOB, AOA, and comammox, and 15 full-length *pmoA* gene sequences from methanotrophs. We limited the database to full-length *amoA* and *pmoA* sequences in order to allow for discovery of primer-binding regions outside of the fragments amplified by conventional *amoA* primer sets (Rotthauwe et al., 1997; Meinhardt et al., 2015). We used the Design Primers option in DECIPHER (Wright et al., 2014) with each one of the candidate

species as the target group and other *amoA/pmoA* sequences entered as closely related groups that should not be amplified. In addition, for all designs, the annealing temperature and PCR conditions were fixed (see Materials and Methods) so that all three species-level primer sets could be simultaneously used in a single thermocycler run. For these species-specific primers, we take advantage of a higher annealing temperature (T_a = 64°C) than what was used for the total comammox (T_a = 48°C) and clade-level comammox (T_a = 52 °C) primers, which is a strategy that also helps improve primer specificity.
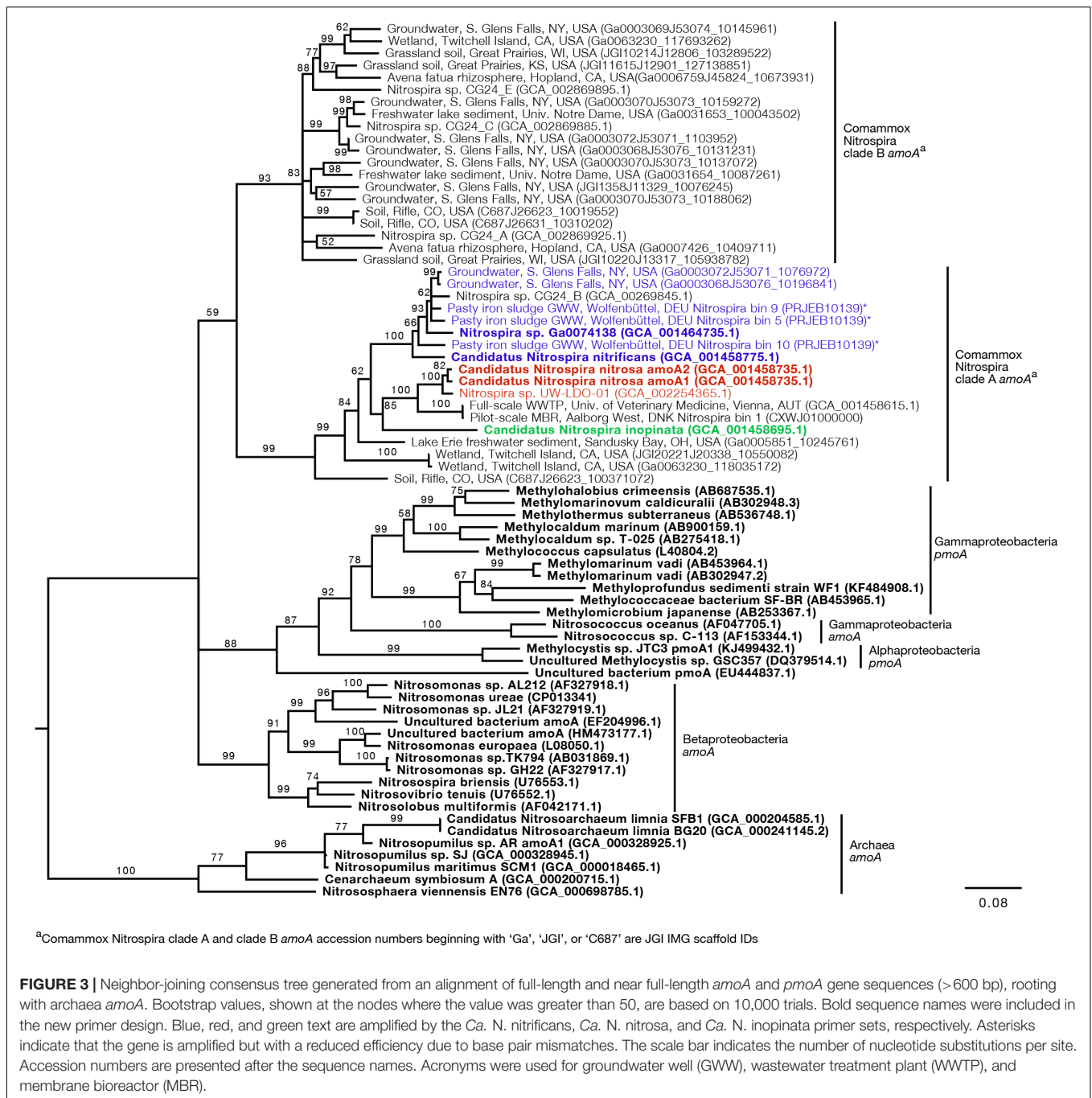
For *Ca.* N. nitrosa, the two *amoA* copies identified in the *Ca.* N. nitrosa genome (van Kessel et al., 2015) were used as the target group. The design algorithm identified primer binding regions at positions 469–493 and 794–812, predicting amplification of a 344 bp fragment (**Figure 1**). The target regions for each primer had 5 or greater mismatches to all other sequences in the design database, predicting a 100% specificity to the target group.

A single sequence (Daims et al., 2015) served as the target group for the design of *Ca.* N. inopinata primers, resulting in primer binding sites at 410–436 and 798–815, for a 406 bp fragment amplification. The target region for each primer had 5 or greater mismatches to other *amoA* and *pmoA* sequences used as non-targets in the design database, also predicting 100% specificity.

For *Ca.* N. nitrificans, two sequences were used as targets; the one in the *Ca.* N. nitrificans genome (van Kessel et al., 2015) and the one from *Nitrospira* sp. Ga0074138 (Pinto et al., 2016), which had 91.4% sequence identity to the *Ca.* N. nitrificans sequence. The primers amplified a 374 bp fragment, between positions 463 and 836. The predicted specificity was again 100%, with 5 or greater mismatches per primer to other sequences used in the design database.

Since the sequence dataset used for design was small, we searched for additional *amoA* sequences that would contain the target regions of the new primer sets. However, this analysis was limited because most of the published *amoA* sequences contain the target side of the forward primers, but not the site targeted by the reverse primer (**Figure 1**). In this search, we found 33 nearly full-length sequences that not only clustered with the comammox sequences but were also long enough to contain the target regions of the species-specific primers (**Figure 3**). This set of nearly full-length *amoA* sequences included *amoA* sequences from comammox metagenomes that have been recently described in the literature (Daims et al., 2015; van Kessel et al., 2015; Palomo et al., 2016; Pinto et al., 2016; Camejo et al., 2017; Pjevac et al., 2017).

An *in silico* primer-target analysis (**Figure 4**) shows that *N. sp.* UW-LDO-01 (Camejo et al., 2017), originally described as a strain of *Ca.* N. nitrosa, has perfect matches to the newly designed primers targeting this species. Two *amoA* sequences (Full scale WWTP, Univ. of Veterinary Medicine, Vienna, AUT, GCA_001458615.1; Pilot-Scale MBR, Aalborg West, DNK Nitrospira bin 1, CXWJ01000000) clustered together and near the *Ca.* N. nitrosa cluster (**Figure 3**). The ANI of their genomes compared to *Ca.* N. nitrosa is 84.5 and 85%, respectively (**Figure 5**), lower than the typical cutoff for species definition (ANI > 94%) (Richter and Rossello-Mora, 2009). In agreement,

FIGURE 3 | Neighbor-joining consensus tree generated from an alignment of full-length and near full-length *amoA* and *pmoA* gene sequences (>600 bp), rooting with archaea *amoA*. Bootstrap values, shown at the nodes where the value was greater than 50, are based on 10,000 trials. Bold sequence names were included in the new primer design. Blue, red, and green text are amplified by the *Ca.* N. nitrificans, *Ca.* N. nitrosa, and *Ca.* N. inopinata primer sets, respectively. Asterisks indicate that the gene is amplified but with a reduced efficiency due to base pair mismatches. The scale bar indicates the number of nucleotide substitutions per site. Accession numbers are presented after the sequence names. Acronyms were used for groundwater well (GWW), wastewater treatment plant (WWTP), and membrane bioreactor (MBR).

we predict that their *amoA* will not be amplified by the nitrosa-specific primer set due to multiple mismatches with forward and reverse primers (**Figure 4**). The *amoA* sequence of *Nitrospira* sp. CG24_B (Palomo et al., 2016) clustered with *Ca.* N. nitrificans (**Figure 3**). However, the ANI of these two genomes (86.3%; **Figure 5**) is below the typical cutoff for species definition and the *in silico* analysis predicts that *Nitrospira* sp. CG24_B *amoA* will not be amplified with the nitrificans-specific primers (**Figure 4**). Three additional *amoA* sequences clustering with *Ca.* N. nitrificans include those from the draft genomes of 'Pasty iron

sludge GWW N. bin 5, N. bin 9, and N. bin 10 (PRJEB10139)' (Daims et al., 2015) (**Figure 3**), which contain fewer mismatches (**Figure 4**) and are predicted to partially amplify (below 56% efficiency) with the *Ca.* N. nitrificans primers despite ANI below the species cutoff (**Figure 5**). Overall, the *amoA* phylogeny close to *Ca.* N. nitrificans remains unresolved, and therefore, the designed primer set for this species will likely require future refinement.

Sequences from clade B comammox were not included in dataset used for primer design, and therefore, an *in silico* analysis

FIGURE 4 | Primer-target mismatch analysis for the newly designed primers and near full-length comammox clade A *amoA* greater than 600 bp. Predicted PCR efficiency is reported separately for the forward and reverse primers in addition to a combined amplification efficiency. Dots indicate a base match and letters indicate a mismatch, with the letter indicating which base is actually present in the sequence.

was also performed with these sequences (**Figure 4** shows *N. sp.* CG24_A *amoA* as an example of clade B comammox). In all cases, the sequences had greater than 4 mismatches per primer and are predicted to not amplify with any of the new species-specific primer sets.
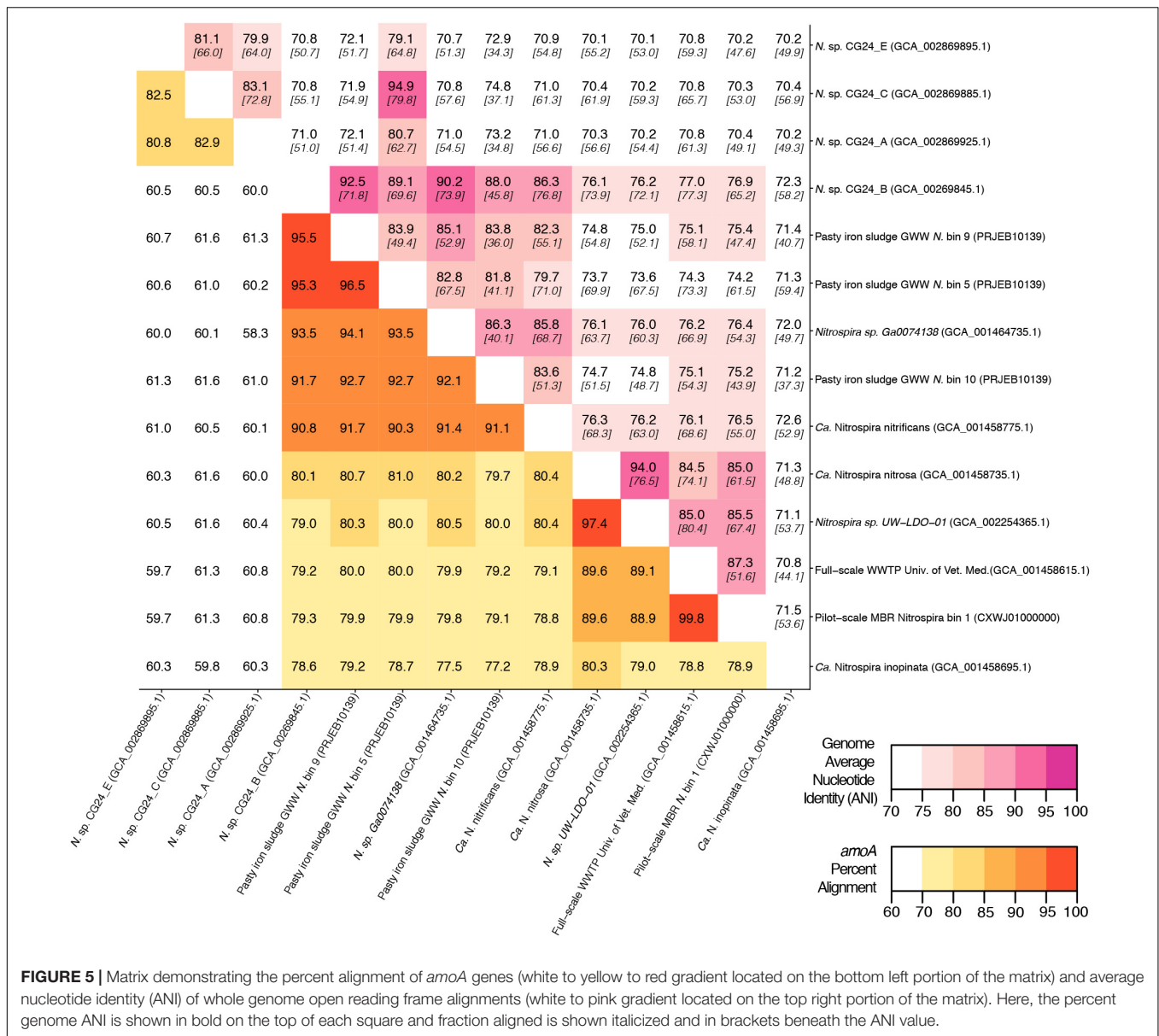
The standard curves for all three primer sets were linear (correlation coefficient $R^2 \geq 0.996$) over 7 orders of magnitude (**Supplementary Table S1**). The amplification efficiency for *Ca.* N. nitrosa, *Ca.* N. inopinata, and *Ca.* N. nitrificans *amoA* was $96 \pm 3.0\%$, $90 \pm 3.0\%$, and $92 \pm 2.4\%$, respectively (**Table 3**). Melting curve analysis of all standards showed amplification of the target product without primer dimer artifacts, represented by a strong fluorescence signal producing a single melting peak at approximately 88.8°C for *Ca.* N. nitrosa (**Supplementary Figures S10A, S11A**), 85.6°C for *Ca.* N. inopinata (**Supplementary Figures S12A, S13A**), and 86.7°C for *Ca.* N. nitrificans (**Supplementary Figures S14A, S15A**).

Environmental Detection of Candidate Comammox Species

Using the new species-specific primer sets developed in this study (Nitrosa amoA-469f/812r, Inopinata amoA-410f/815r,

and Nitrificans amoA463f/836r), we evaluated the samples originating from the low DO BNR bioreactors (**Table 1**). With the new primers, positive detections obtained with melting curve analysis (**Supplementary Figures S10–S15**) correlated well with a positive target amplicon in the agarose gel (**Supplementary Figures S16–S18**). Additionally, cross-hybridization and primer dimers were not observed following PCR with agarose gel electrophoresis with the standard templates (**Supplementary Figure S19**). The results from the new species-specific primers are also summarized in decision matrices that were used to consistently evaluate qPCR results between assays (**Supplementary Tables S6–S8**).

Comammox *amoA* belonging to *Ca.* N. nitrosa were detected in all environmental samples. Agarose gel electrophoresis of the qPCR amplified products validated these results, since a single amplicon with the expected length (344-bp) was obtained from all 19 samples (**Supplementary Figure S17**). *Ca.* N. nitrosa abundance was greater than 10^3 copies *amoA*/ng DNA in nine samples originating from low DO bioreactors L_SBR, P_SBR, P_CF, and TRA (**Figure 2D**). The maximum number of *amoA* copies was obtained from L_SBR-1, with approximately 2.5×10^4 copies *amoA*/ng DNA (**Figure 2D**). After 45 cycles, six samples (L_SBR-4, P_CF-4, NS-1, NS-2, NS-3, and NS-4) were estimated

	E	C	A	B	bin9	bin5	Ga0074138	bin10	nitrificans	nitrosa	UW-LDO-01	Vet.Med.	MBR bin1	inopinata	
E		81.1 [66.0]	79.9 [64.0]	70.8 [50.7]	72.1 [51.7]	79.1 [64.8]	70.7 [51.3]	72.9 [34.3]	70.9 [54.8]	70.1 [55.2]	70.1 [53.0]	70.8 [59.3]	70.2 [47.6]	70.2 [49.9]	N. sp. CG24_E (GCA_002869895.1)
C	82.5		83.1 [72.8]	70.8 [55.1]	71.9 [54.9]	94.9 [79.8]	70.8 [57.6]	74.8 [37.1]	71.0 [61.3]	70.4 [61.9]	70.2 [59.3]	70.8 [65.7]	70.3 [53.0]	70.4 [56.9]	N. sp. CG24_C (GCA_002869885.1)
A	80.8	82.9		71.0 [51.0]	72.1 [51.4]	80.7 [62.7]	71.0 [54.5]	73.2 [34.8]	71.0 [56.6]	70.3 [56.6]	70.2 [54.4]	70.8 [61.3]	70.4 [49.1]	70.2 [49.3]	N. sp. CG24_A (GCA_002869925.1)
B	60.5	60.5	60.0		92.5 [71.8]	89.1 [69.6]	90.2 [73.9]	88.0 [45.8]	86.3 [76.8]	76.1 [73.9]	76.2 [72.1]	77.0 [77.3]	76.9 [65.2]	72.3 [58.2]	N. sp. CG24_B (GCA_00269845.1)
bin9	60.7	61.6	61.3	95.5		83.9 [49.4]	85.1 [52.9]	83.8 [36.0]	82.3 [55.1]	74.8 [54.8]	75.0 [52.1]	75.1 [58.1]	75.4 [47.4]	71.4 [40.7]	Pasty iron sludge GWW N. bin 9 (PRJEB10139)
bin5	60.6	61.0	60.2	95.3	96.5		82.8 [67.5]	81.8 [41.1]	79.7 [71.0]	73.7 [69.9]	73.6 [67.5]	74.3 [73.3]	74.2 [61.5]	71.3 [59.4]	Pasty iron sludge GWW N. bin 5 (PRJEB10139)
Ga0074138	60.0	60.1	58.3	93.5	94.1	93.5		86.3 [40.1]	85.8 [68.7]	76.1 [63.7]	76.0 [60.3]	76.2 [66.9]	76.4 [54.3]	72.0 [49.7]	Nitrospira sp. Ga0074138 (GCA_001464735.1)
bin10	61.3	61.6	61.0	91.7	92.7	92.7	92.1		83.6 [51.3]	74.7 [51.5]	74.8 [48.7]	75.1 [54.3]	75.2 [43.9]	71.2 [37.3]	Pasty iron sludge GWW N. bin 10 (PRJEB10139)
nitrificans	61.0	60.5	60.1	90.8	91.7	90.3	91.4	91.1		76.3 [68.3]	76.2 [63.0]	76.1 [68.6]	76.5 [55.0]	72.6 [52.9]	Ca. Nitrospira nitrificans (GCA_001458775.1)
nitrosa	60.3	61.6	60.0	80.1	80.7	81.0	80.2	79.7	80.4		94.0 [76.5]	84.5 [74.1]	85.0 [61.5]	71.3 [48.8]	Ca. Nitrospira nitrosa (GCA_001458735.1)
UW-LDO-01	60.5	61.6	60.4	79.0	80.3	80.0	80.5	80.0	80.4	97.4		85.0 [80.4]	85.5 [67.4]	71.1 [53.7]	Nitrospira sp. UW–LDO–01 (GCA_002254365.1)
Vet.Med.	59.7	61.3	60.8	79.2	80.0	80.0	79.9	79.2	79.1	89.6	89.1		87.3 [51.6]	70.8 [44.1]	Full-scale WWTP Univ. of Vet. Med.(GCA_001458615.1)
MBR bin1	59.7	61.3	60.8	79.3	79.9	79.9	79.8	79.1	78.8	89.6	88.9	99.8		71.5 [53.6]	Pilot–scale MBR Nitrospira bin 1 (CXWJ01000000)
inopinata	60.3	59.8	60.3	78.6	79.2	78.7	77.5	77.2	78.9	80.3	79.0	78.8	78.9		Ca. Nitrospira inopinata (GCA_001458695.1)

Genome Average Nucleotide Identity (ANI) 70 75 80 85 90 95 100

amoA Percent Alignment 60 70 80 85 90 95 100

FIGURE 5 | Matrix demonstrating the percent alignment of *amoA* genes (white to yellow to red gradient located on the bottom left portion of the matrix) and average nucleotide identity (ANI) of whole genome open reading frame alignments (white to pink gradient located on the top right portion of the matrix). Here, the percent genome ANI is shown in bold on the top of each square and fraction aligned is shown italicized and in brackets beneath the ANI value.

to have fewer copies of *amoA* per ng DNA than the calculated LOD (*M* = 11, 95% CI [9.3, 13] copies/ng DNA) (**Figure 2D** and **Table 3**). However, these positive qPCR detections were confirmed via agarose gel electrophoresis (**Supplementary Figures S16A, S17**), suggesting that the *Ca.* N. nitrosa primers are highly sensitive (**Supplementary Table S6**). The melting curve analysis at 45 cycles (**Supplementary Figure S10**) and 30 cycles (**Supplementary Figure S11**) also confirmed the positive qPCR detections.

Compared to the quantification of total comammox (Fowler et al., 2017), *Ca.* N. nitrosa comprised an average of 340 ± 70%, 96 ± 29%, 67 ± 13%, and 93% of the total comammox population in the true positive detections obtained from L_SBR, P_SBR, P_CF, and TRA, respectively (**Figures 2A,D**). Although in some cases the lower quantification with the new species-specific

primers could indicate the presence of other non-targeted comammox species, we also see that the total comammox primers contain 4 total mismatches to the primer-target region of *Ca.* N. nitrosa, since mismatches were allowed in the original total comammox primer design (**Supplementary Table S2**). In addition, we observed reduced total comammox (Fowler et al., 2017) efficiency when *Ca.* N. nitrosa *amoA* was present at less than 10^3 copies (**Supplementary Table S1**). Thus, total comammox (Fowler et al., 2017) primer mismatches to *Ca.* N. nitrosa *amoA* may contribute to an overall reduced amplification efficiency and underestimation of this species in environmental samples that contain an abundant comammox population, like L_SBR and P_SBR. All in all, the new species-specific primers show that *Ca.* N. nitrosa is an important member of the comammox population in wastewater treatment plants and may

contribute to low oxygen nitrification (as seen in L_SBR, P_SBR, P_CF and TRA).

Comammox *amoA* belonging to *Ca.* N. inopinata and *Ca.* N. nitrificans were less frequent and were only detected with less than 10^2 copies *amoA*/ng DNA in all samples (**Figures 2E,F** and **Supplementary Figure S18**). N. inopinata *amoA* was positively detected in one sample from the Madison Metropolitan Sewerage District Nine Springs plant (NS), NS-1 (**Supplementary Table S7**); although, the concentration of N. inopinata *amoA* was less than the calculated LOD (**Figure 2E** and **Table 3**). *Ca.* N. nitrificans was only detected above the LOQ in L_SBR-1 and L_SBR-2 with approximately 86 and 21 *amoA*/ng DNA, respectively (**Figure 2F** and **Supplementary Table S8**). Overall, comammox *amoA* belonging to *Ca.* N. inopinata or *Ca.* N. nitrificans were minor contributors to the total bacteria population after normalization (with relative abundances less than 0.03%) (**Supplementary Figure S20**).

With the new species-specific primers, we intentionally ran amplifications for 45 cycles, to make sure we could detect any potential non-target amplification. Consequently, the melting curve analysis show some peaks in samples that did not have the target (**Supplementary Figures S12B–F, S14C–F**), but these peaks corresponds to noise after a large number of cycles that would not be seen in quantifications with a more conventional number of cycles. To verify that the noise in the melting peaks was due to the large number of cycles, we repeated the qPCR experiment with just 30 cycles, which eliminated the peaks that corresponded to noise (**Supplementary Figures S11, S13, S15**). Agarose gel electrophoresis of the qPCR amplified products after 30 cycles reaffirm the specificity of the new species-specific primers (**Supplementary Figure S16**). Moreover, samples that contained noise in the melting peaks when the qPCR was performed with 45 cycles occurred at a late C_q, and consequently, any artifact associated fluorescence was minimal. If these samples had been quantified, the concentrations would be below both the LOD and LOQ, making them easy to eliminate from further analysis without needing to confirm with agarose gel electrophoresis (**Supplementary Tables S6– S8**).

DISCUSSION

Since the late 1800s, nitrification had been described as a division of labor between two phylogenetically distinct and specialized chemolithoautotrophs, the ammonia oxidizers and the nitrite oxidizers (Brock, 1999). This perception was not challenged for over a century, until Costa et al. (2006) hypothesized that a single nitrifying bacterium combining ammonia oxidation and nitrite oxidation should exist in nature. Nearly one decade later, in late 2015, two separate research groups discovered, cultivated, and characterized the 'missing' comammox organism from an aquaculture system and a deep oil exploration well, respectively (Daims et al., 2015; van Kessel et al., 2015). To date, there is still uncertainty about the occurrence of the novel *Nitrospira*-like comammox organisms in full scale activated sludge systems. Some studies suggest

that comammox are not relevant to conventional wastewater treatment (Chao et al., 2016; Gonzalez-Martinez et al., 2016). However, since comammox were discovered in low oxygen environments (Daims et al., 2015; van Kessel et al., 2015), they may be able to efficiently utilize oxygen and may be important to low DO BNR, as recently described by Camejo et al. (2017).

Long-term low DO total nitrogen removal has been successfully demonstrated in laboratory (Camejo et al., 2016) and pilot-scale (Camejo et al., 2016; Keene et al., 2017) bioreactors seeded with Nine Springs WWTP activated sludge and operated with DO concentrations less than 0.60 mg O_2/L. Samples analyzed from the Nine Springs WWTP showed presence of *Ca.* N. nitrosa, but below the LOQ. The sludge from this WWTP was the seed for the L_SBR, P_SBR, and P_CF reactors, all of which showed presence of *Ca.* N. nitrosa, suggesting that the low DO conditions favored accumulation of *Ca.* N. nitrosa as a potential participant in ammonia oxidation in low DO BNR. Interestingly, *Ca.* N. nitrosa was also abundant in the full-scale TRA system, which is operated with low DO conditions (**Table 1**), providing independent support for the hypothesis that *Ca.* N. nitrosa is an important contributor to ammonia oxidation in WWTP operated with low DO. *Ca.* N. nitrificans was detected above the LOQ in the L_SBR reactor only, whereas *Ca.* N. inopinata was not detected above the LOD in any of the low DO reactors; therefore, out of the three candidate species, *Ca.* N. nitrosa appears to be the only one that becomes enriched under low DO conditions.

When compared to quantification of the total bacterial population by 16S rRNA-targeted qPCR, the greatest relative abundance of *Ca.* N. nitrosa was 10% and 6% of total bacteria in samples originating from the laboratory-scale and pilot-scale sequencing batch reactors, respectively (**Supplementary Figure S20**). Although these two reactors were seeded with the same sludge, they were fed from completely different sources (synthetic media for L_SBR and full-scale primary effluent for P_SBR), but operated with low oxygen and a long solids retention time (SRT, both 80 days) (Camejo et al., 2016). The continuous flow reactors (P_CF and TRA) had lower relative abundances of *Ca.* N. nitrosa, and although operated with low DO, they had a much shorter SRT (both 10 days; **Table 1**). Thus, in addition to oxygen, SRT may be a factor that contributes to comammox abundance in low DO reactors.

Detection and quantification of microorganisms via real-time PCR relies on designing primers with good specificity to the targeted organisms, good coverage of the targeted group, and good quality of the experimental results. With any newly discovered target group, the quality of primer sets to achieve accurate quantification depends on the quality of the databases used for design and the design considerations. As more sequences of the targeted organisms become available, designs can only improve. The first primer design for comammox quantification described sets for targeting clade A and B *amoA* within the *Nitrospira* genus (Pjevac et al., 2017), and subsequently another set of primers was published, aiming for greater coverage to target all comammox *amoA* (Fowler et al., 2017). Our evaluation of

these primer sets in environmental samples from low DO BNR plants showed challenges with primer dimer formation when target sequences were not abundant (clade A and clade B primers, Pjevac et al., 2017), unspecific amplifications in environmental samples (all sets), and underestimation with some strains due to primer-target mismatches (total comammox, Fowler et al., 2017). These challenges are more important when designing for broad target groups. In general, melting curve analyses of qPCR results are helpful for detection of nonspecific PCR products since different fragments will typically appear as distinct melting peaks (Ririe et al., 1997; Wittwer and Kusakawa, 2004). However, the broad detection and quantification of comammox amoA has high GC content variability among the comammox organisms described thus far (**Table 3**). Therefore, a wide range of amplicon melting temperatures are expected when using the total comammox (Fowler et al., 2017) or clade-level (Pjevac et al., 2017) assays, making melting curve analyses difficult to interpret. We found that verifying amplified products with the total comammox (Fowler et al., 2017) or clade-level (Pjevac et al., 2017) primers cannot be completed with melting curve analysis alone and will require a subsequent agarose gel electrophoresis of our environmental samples, which is not ideal for routine qPCR applications.

Setting the challenge of broad comammox quantification aside, we aimed at designing primer sets specific to the three candidate species described thus far. Evidently, there is a very small number of sequences representative of these candidate species; therefore, the designed primers inherently have 100% coverage. Specificity depends on finding sufficient differences between target and non-targets sequences (Wright et al., 2014), and the results showed this was possible for the three candidate species targeted. As more sequences become available from metagenome assemblies, the specificity of these primers can be re-evaluated. The beginning of such activity was performed in this study (**Figure 4**). Importantly, the qPCR experiments with the newly designed primers were conducted with a longer number of thermal cycles than typical (45 cycles) to increase the chances of detecting potential non-specific amplifications. This extended thermal cycling confirmed that that the primers are highly specific and are not producing unwanted non-specific amplifications.

Taken together, the species-specific comammox primers designed in this study enabled an analysis of comammox abundance focused solely on the candidate comammox species currently described in the literature. Although a narrowly focused analysis by design, it eliminated the problems of unspecific amplifications that compromised the use of broad comammox primers and resulted in strong experimental evidence in support of *Ca.* N. nitrosa being a comammox organism in energy-efficient BNR systems operated under low DO conditions. Given earlier studies on low-DO BNR providing some evidence of AOB also present in these systems (Fitzgerald et al., 2015; Keene et al., 2017) and no evidence of AOA (Fitzgerald et al., 2015), we hypothesize that *Ca.* N. nitrosa has an important contribution to ammonia oxidation in low-DO BNR reactors.

DATA AVAILABILITY STATEMENT

All datasets analyzed in this study are included in the manuscript and the **Supplementary Files**.

AUTHOR CONTRIBUTIONS

NB and DN developed the research plan and project goals. NB designed the primers and performed the laboratory work. Both NB and DN drafted the manuscript, tables, and figures. DN contributed sources of project funding. All authors have critically read, corrected, and approved the final version of the manuscript and agreed with the opinions expressed here.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.00036/full#supplementary-material

REFERENCES

Armbruster, D. A., and Pry, T. (2008). Limit of blank, limit of detection and limit of quantitation. *Clin. Biochem. Rev.* 29(Suppl. 1), S49–S52.

Bartelme, R. P., McLellan, S. L., and Newton, R. J. (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. doi: 10.3389/fmicb.2017.00101

Brock, T. D. (1999). *Milestones in Microbiology 1546 to 1940*. Washington, D.C: American Society Mic Series: ASM Press.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55, 611–622. doi: 10.1373/clinchem.2008.112797

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Camejo, P. Y., Owen, B. R., Martirano, J., Ma, J., Kapoor, V., Santo Domingo, J., et al. (2016). Candidatus *Accumulibacter phosphatis* clades enriched under cyclic anaerobic and microaerobic conditions simultaneously use different electron acceptors. *Water Res.* 102, 125–137. doi: 10.1016/j.watres.2016.06.033

Camejo, P. Y., Santo Domingo, J., McMahon, K. D., and Noguera, D. R. (2017). Genome-Enabled insights into the ecophysiology of the comammox bacterium "candidatus nitrospira nitrosa. *mSystems* 2, e0059-17. doi: 10.1128/mSystems.00059-17

Chao, Y., Mao, Y., Yu, K., and 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. doi: 10.1007/s00253-016-7655-9

Costa, E., Perez, J., and Kreft, J. U. (2006). Why is metabolic labour divided in nitrification? *Trends Microbiol.* 14, 213–219. doi: 10.1016/j.tim.2006.03.006

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. doi: 10.1038/nature16461

Fitzgerald, C. M., Camejo, P., Oshlag, J. Z., and Noguera, D. R. (2015). Ammonia-oxidizing microbial communities in reactors with efficient nitrification at low-dissolved oxygen. *Water Res.* 70, 38–51. doi: 10.1016/j.watres.2014.11.041

Fowler, S. J., Palomo, A., Dechesne, A., Mines, P. D., and Smets, B. F. (2017). Comammox nitrospira are abundant ammonia oxidizers in diverse groundwater-fed rapid sand filter communities. *Environ. Microbiol.* 20, 1002–1015. doi: 10.1111/1462-2920.14033

Gonzalez-Martinez, A., Rodriguez-Sanchez, A., van Loosdrecht, M. C. M., Gonzalez-Lopez, J., and Vahala, R. (2016). Detection of comammox bacteria in full-scale wastewater treatment bioreactors using tag-454-pyrosequencing. *Environ. Sci. Pollut. Res. Int.* 23, 25501–25511. doi: 10.1007/s11356-016-7914-4

Hu, H. W., and He, J. Z. (2017). Comammox—a newly discovered nitrification process in the terrestrial nitrogen cycle. *J. Soils Sedim.* 17, 2709–2717. doi: 10.1007/s11368-017-1851-9

Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119

Karst, S. M., Kirkegaard, R. H., and Albertsen, M. (2016). mmgenome: a toolbox for reproducible genome extraction from metagenomes. *bioRxiv* [Preprint]. doi: 10.1101/059121

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Keene, N. A., Reusser, S. R., Scarborough, M. J., Grooms, A. L., Seib, M., Santo Domingo, J., et al. (2017). Pilot plant demonstration of stable and efficient high rate biological nutrient removal with low dissolved oxygen conditions. *Water Res.* 121(Suppl. C), 72–85. doi: 10.1016/j.watres.2017.05.029

Keene-Beach, N., and Noguera, D. R. (2018). Design and assessment of species-level qPCR primers targeting comammox. *bioRxiv* [Preprint]. doi: 10.1101/348664

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., and Horn, M. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 41:e1. doi: 10.1093/nar/gks808

Koops, H. P., Purkhold, U., Pommerening-Röser, A., Timmermann, G., and Wagner, M. (2006). "The lithoautotrophic ammonia-oxidizing bacteria," in *The Prokaryotes*, Vol. 5, eds M. Dworkin, S. Falkow, E. Rosenberg, K. H. Schleifer, and E. Stackebrandt (New York, NY: Springer-Verlag), 778–811. doi: 10.1007/0-387-30745-1-36

Markowitz, V. M., Chen, I. M., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., et al. (2012). IMG: the integrated microbial genomes database and comparative

analysis system. *Nucleic Acids Res.* 40, D115–D122. doi: 10.1093/nar/gkr1044

Meinhardt, K. A., Bertagnolli, A., Pannu, M. W., Strand, S. E., Brown, S. L., and Stahl, D. A. (2015). Evaluation of revised polymerase chain reaction primers for more inclusive quantification of ammonia-oxidizing archaea and bacteria. *Environ. Microbiol. Rep.* 7, 354–363. doi: 10.1111/1758-2229.12259

NCBI Resource Coordinators (2017). Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 45, D12–D17. doi: 10.1093/nar/gkw1071

Palomo, A., Jane Fowler, S., Gulay, A., Rasmussen, S., Sicheritz-Ponten, T., and Smets, B. F. (2016). Metagenomic analysis of rapid gravity sand filter microbial communities suggests novel physiology of *Nitrospira* spp. *ISME J.* 10, 2569–2581. doi: 10.1038/ismej.2016.63

Palomo, A., Pedersen, A. G., Fowler, S. J., Dechesne, A., Sicheritz-Ponten, T., and Smets, B. F. (2018). Comparative genomics sheds light on niche differentiation and the evolutionary history of comammox *Nitrospira*. *ISME J.* 12, 1779–1793. doi: 10.1038/s41396-018-0083-3

Paredes, D., Kuschk, P., Mbwette, T. S. A., Stange, F., Müller, R. A., and Köser, H. (2007). New aspects of microbial nitrogen transformations in the context of wastewater treatment – a review. *Eng. Life Sci.* 7, 13–25. doi: 10.1002/elsc.200620170

Park, H. D., and Noguera, D. R. (2004). Evaluating the effect of dissolved oxygen on ammonia-oxidizing bacterial communities in activated sludge. *Water Res.* 38, 3275–3286. doi: 10.1016/j.watres.2004.04.047

Pester, M., Rattei, T., Flechl, S., Grongroft, A., Richter, A., Overmann, J., et al. (2012). amoA-based consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of amoA genes from soils of four different geographic regions. *Environ. Microbiol.* 14, 525–539. doi: 10.1111/j.1462-2920.2011.02666.x

Pinto, A. J., Marcus, D. N., Ijaz, U. Z., Bautista-de Lose Santos, Q. M., Dick, G. J., and Raskin, L. (2016). Metagenomic evidence for the presence of comammox nitrospira-like bacteria in a drinking water system. *mSphere* 1:e0054-15. doi: 10.1128/mSphere.00054-15

Pjevac, P., Schauberger, C., Poghosyan, L., Herbold, C. W., van Kessel, M., 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. doi: 10.3389/fmicb.2017.01508

Purkhold, U., Pommerening-Roser, A., Juretschko, S., Schmid, M. C., Koops, H. P., and 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. doi: 10.1128/aem.66.12.5368-5382.2000

Richter, M., and Rossello-Mora, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19126–19131. doi: 10.1073/pnas.0906412106

Richter, M., Rossello-Mora, R., Oliver Glockner, F., and Peplies, J. (2016). JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32, 929–931. doi: 10.1093/bioinformatics/btv681

Ririe, K. M., Rasmussen, R. P., and Wittwer, C. T. (1997). Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Anal. Biochem.* 245, 154–160. doi: 10.1006/abio.1996.9916

Rotthauwe, J. H., Witzel, K. P., and 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.

Stadler, L. B., and Love, N. G. (2016). Impact of microbial physiology and microbial community structure on pharmaceutical fate driven by dissolved oxygen concentration in nitrifying bioreactors. *Water Res.* 104, 189–199. doi: 10.1016/j.watres.2016.08.001

Thijs, S., Op De Beeck, M., Beckers, B., Truyens, S., Stevens, V., Van Hamme, J. D., et al. (2017). Comparative evaluation of four bacteria-specific primer pairs for 16S rRNA gene surveys. *Front. Microbiol.* 8:494. doi: 10.3389/fmicb.2017.00494

van Kessel, M. A., Speth, D. R., Albertsen, M., Nielsen, P. H., Op den Camp, H. J., Kartal, B., et al. (2015). Complete nitrification by a single microorganism. *Nature* 528, 555–559. doi: 10.1038/nature16459

Wittwer, C. T., and Kusakawa, N. (2004). "Real-time PCR," in *Molecular Microbiology: Diagnostic Principles and Practice*, eds D. H. Persing, F. C.

Tenover, R. T. Hayden, and A. Van Belkum (Washington, DC: ASM Press).

Wright, E. S. (2015). DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics* 16:322. doi: 10.1186/s12859-015-0749-z

Wright, E. S. (2016). Using DECIPHER v2.0 to analyze big biological sequence data in R. *R J.* 8, 352–359. doi: 10.18129/B9.bioc.DECIPHER

Wright, E. S., Yilmaz, L. S., Ram, S., Gasser, J. M., Harrington, G. W., and Noguera, D. R. (2014). Exploiting extension bias in polymerase chain reaction to improve primer specificity in ensembles of nearly identical DNA templates. *Environ. Microbiol.* 16, 1354–1365. doi: 10.1111/1462-2920.12259