



Comammox Nitrospira are abundant ammonia oxidizers in diverse groundwaterfed rapid sand filter communities

Fowler, Susan Jane; Palomo, Alejandro; Dechesne, Arnaud; Mines, Paul D.; Smets, Barth F.

Published in:
Environmental Microbiology

Link to article, DOI:
[10.1111/1462-2920.14033](https://doi.org/10.1111/1462-2920.14033)

Publication date:
2018

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):

Fowler, S. J., Palomo, A., Dechesne, A., Mines, P. D., & Smets, B. F. (2018). Comammox Nitrospira are abundant ammonia oxidizers in diverse groundwaterfed rapid sand filter communities. *Environmental Microbiology*, 20(3), 1002-1015. <https://doi.org/10.1111/1462-2920.14033>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Comammox *Nitrospira* are abundant ammonia oxidizers in diverse groundwater-fed rapid sand filter communities

5 S. Jane Fowler¹, Alejandro Palomo¹, Arnaud Dechesne¹, Paul D. Mines², Barth F. Smets^{1*}

¹Department of Environmental Engineering, ²Department of Micro- and Nanotechnology, Technical University of Denmark, Kgs Lyngby 2800, Denmark.

10 **Running title:** Comammox *Nitrospira* in drinking water biofilters

*Address correspondence to Barth F. Smets: Bygningstorvet, Building 115, Kongens Lyngby, Denmark 2800 email: b fsm@env.dtu.dk, phone: +45 45251600.

Significance Statement

Determining the contribution of the recently discovered completely nitrifying (comammox) *Nitrospira* to global nitrogen cycling necessitates tools for their detection and quantification in diverse environments. Herein, we present an assay which enables the simultaneous detection and quantification of a broad range of clades A and B comammox *Nitrospira*, a key step in understanding their abundance, diversity and distribution in natural and engineered environments. We apply this method to groundwater-fed rapid sand filters at 12 drinking water treatment plants, and examine the nitrifying communities in these biological filters. Our results demonstrate that comammox *Nitrospira* are dominant nitrifiers in groundwater-fed biofilters.

Summary

The recent discovery of completely nitrifying *Nitrospira* demands a re-examination of nitrifying environments to evaluate their contribution to nitrogen cycling. To approach this challenge, tools are needed to detect and quantify comammox *Nitrospira*. We present primers for the simultaneous quantification and diversity assessment of both comammox *Nitrospira* clades. The primers cover a wide range of comammox diversity, spanning all available high quality sequences. We applied these primers to 12 groundwater-fed rapid sand filters, and found comammox *Nitrospira* to be abundant in all filters. Clade B comammox comprise the majority (~75%) of comammox abundance in all filters. Nitrosomonadaceae were present in all filters, though at low abundance (mean=1.8%). Ordination suggests that temperature impacts the structure of nitrifying communities, and in

particular that increasing temperature favours *Nitrospira*. The nitrogen content of the filter
40 material, sulfate concentration, and surface ammonium loading rates shape the structure of
the comammox guild in the filters. This work provides an assay for simultaneous detection
and diversity assessment of clade A and B comammox *Nitrospira*, expands our current
knowledge of comammox *Nitrospira* diversity, and demonstrates a key role for comammox
Nitrospira in nitrification in groundwater-fed biofilters.

45

Introduction

The conventional understanding of nitrification since its discovery in the 1890s is that
it is a two-step process carried out by distinct groups of chemolithoautotrophic nitrifiers:
Ammonia oxidizers and nitrite oxidizers. The repertoire of nitrifiers was expanded in 2005
50 with the discovery of ammonia oxidizing archaea (AOA)(Könneke *et al.*, 2005), but this still
did not challenge the prevailing perception that labour was divided in nitrification. Though
the reasons for this observed division of labour were unclear, it was predicted that a
complete nitrifier could theoretically exist, and would thrive in surface attached
environments with low ammonium loading (Costa *et al.*, 2006). The simultaneous discovery
55 of comammox *Nitrospira* by a number of groups (Daims *et al.*, 2015; Pinto *et al.*, 2015; van
Kessel *et al.*, 2015; Palomo *et al.*, 2016) disrupts our long-held understanding of nitrification.
The low prevalence of comammox *Nitrospira* in most environments examined to date
(Pjevac *et al.*, 2017), combined with their resistance to cultivation, explains how these
organisms were overlooked for so long. The sole enrichments of comammox bacteria come
60 from environments where they made up relatively small fractions of the original
communities (Daims *et al.*, 2015; van Kessel *et al.*, 2015). This begs the question: Do
comammox *Nitrospira* play a significant role in nitrogen cycling, or are they typically found

only in low abundance in diverse environments? We know very little about the ecology of comammox bacteria and their global contribution to ammonia oxidation. To begin to
65 investigate these questions, tools to detect and quantify comammox *Nitrospira* are required.

Evidence from metagenomic sequencing suggests that comammox *Nitrospira* are one of the most dominant taxa at Islevbro waterworks in Zealand, Denmark (Palomo *et al.*, 2016). Amplicon-based sequencing of the 16S rRNA gene in additional groundwater-fed rapid sand filters (RSF) reveal that *Nitrospira* spp. are abundant in many examined RSFs, though
70 whether these include comammox is heretofore unknown (Albers *et al.*, 2015; Gülay *et al.*, 2016). Interestingly, these filters represent exactly the type of environment which Costa *et al.* (2006) predicted comammox would thrive in, namely, surface attached communities exposed to low ammonium loading. In this work, we examined the microbial communities of a range of RSFs and report on the abundance, diversity and composition of nitrifiers,
75 including comammox *Nitrospira*, using qPCR and amplicon sequencing. We present a new high coverage qPCR assay for the *amoA* gene of *Nitrospira* - the first to simultaneously target both comammox *Nitrospira* clades A and B as verified by clone library and amplicon sequencing. Lastly, to gain insight into the ecology of comammox *Nitrospira*, we assessed the contribution of physicochemical and operational parameters of the filters to the
80 composition of the nitrifying communities and the comammox *Nitrospira* guild, and examined the ratios of nitrifying guilds to identify relationships between nitrifiers. As the filters and source aquifers share similar characteristics with previously examined filters, we hypothesized that total *Nitrospira* and comammox *Nitrospira* would be abundant in all filters, but that variations in influent loading characteristics and physicochemical features of the
85 filter environment would play a role in shaping the *Nitrospira* diversity in individual filters.

Results

Amplification of comammox amoA

Primers for the amplification and quantification of comammox *amoA* clades A and B were designed based on *amoA* genes from enriched or high quality metagenome-derived genomes of comammox organisms. Clone library sequencing resulted in the recovery of 19 unique sequences (from 21 clones), corresponding to 14 clade A and 5 clade B comammox *amoA* sequences. Comammox *amoA* amplicons from 12 waterworks were then sequenced by Illumina MiSeq, resulting in 88 unique *amoA* gene fragments, 41 from clade A, and 47 from clade B. When applying these primers to the quantification of *Nitrospira amoA* genes, the standard curve of the reaction was linear between 10^1 - 10^9 copies/reaction, so as few as 10 copies can theoretically be quantified (Figure S1). However, inspection of melt curves (Figure S2) showed that noise existed in samples with low copy numbers and that reliable quantification of between 100-1000 or more copies is possible in environmental samples. Although this points towards non-specific amplification, additional bands were never observed when amplicons were run on an agarose gel (Figure S3). Primer coverage was evaluated *in silico* using a database of 48 *amoA Nitrospira* sequences (not including clones from this study) of the correct gene region obtained from isolates or enriched strains, metagenome-derived genomes and metagenomes (Figure S4). When allowing for up to two mismatches, both primers covered all sequences. In the analysis of clone sequences only, all but one of the clone sequences were amplified when two mismatches were allowed, suggesting that coverage estimates from *in silico* analysis are conservative as *in situ* amplification can extend beyond two mismatches.

110 *Filter communities*

The nitrifying communities of 12 waterworks were characterized by sequencing of 16S rRNA Bacteria, *nxB Nitrospira* and *amoA Nitrospira* gene amplicons as well as by qPCR-based quantification of 16S rRNA genes of Bacteria, AOB, and *Nitrospira*, *nxB* of *Nitrospira* and *Nitrobacter* and *amoA* of *Nitrospira* and AOA (Table 1). Though variations were observed even between biological replicates of the same filter, as observed previously (Gülay *et al.*, 2016), the microbial communities of all filters shared similar compositions. Total cell numbers varied substantially between filters, from 4.5×10^7 - 1.2×10^{10} /g of filter material (median 2.6×10^9 /g). The after-filter at DWTP-7 contained unusually low cell numbers (4.5×10^7 /g), which is likely because the majority of biological activity occurs in the long (2.3 m) pre-filter at this waterworks. The remaining filters had much higher total cell numbers (Figure 1A). Nitrifiers made up between 24- 59% of the communities in all filters with the exception of DWTP-11 pre-filter (5.2%) (Figure 1B). Pre-filters are commonly believed to be mainly the site of abiotic iron oxidation and precipitation (Tatari *et al.*, 2017), and this appears to be corroborated based on the relatively lower density of nitrifiers in DWTP-11. *Nitrospira* spp. represented the vast majority of nitrifying bacteria in the filter communities with 112 ASVs, and made up an average of 34% of filter communities. This high abundance is reflected in both qPCR and sequence data (Figure 1AB). In select filters, amplicon sequencing and qPCR indicated that *Nitrospira* spp. comprised a smaller proportion of the community (DWTP-9 (16.1%), and DWTP-11 pre-filter (4.8%)), but in the remaining filters, *Nitrospira* made up between 24.2% (DWTP-12) and 58.8% (DWTP-2) of the microbial communities. Quantification of *Nitrospira* spp. by qPCR of the 16S rRNA and *nxB* genes generally resulted in similar estimates of *Nitrospira* abundance, although the ratios within samples varied somewhat, which may be due to variation in coverage of the primer sets and the copy number of *nxB* in the *Nitrospira* present in different filters (Figure 1A). We assumed a

135 genomic copy number of one *nxB* per *Nitrospira* spp., as metagenomic analysis revealed that the dominant *Nitrospira* spp. at DWTP-8 each harboured a single *nxB* gene (Palomo et al. 2017), however some *Nitrospira* spp. can harbour up to six copies of *nxB* (Pester et al., 2014). Relative to qPCR of *nxB* and *Nitrospira* 16S rRNA, Illumina sequencing of universal Bacteria 16S rRNA gene amplicons tended to overestimate the abundance of *Nitrospira* (Figure 1C). AOB from the family Nitrosomonadaceae were present in all filters, though generally comprised less than 2.5% of the communities according to both qPCR and sequence data (Figure 1AD). Of 60 ASVs, 33 were classified as *Nitrosomonas* or *Nitrospira* spp. while the others were not classified below the family level. Nitrosomonadaceae abundance was particularly high in DWTP-9 (7.8%), and DWTP-10 (7.3%). Interestingly, DWTP-9 was a filter in which *Nitrospira* abundance was considerably lower relative to other filters, suggesting that AOB may play a larger role in ammonia oxidation in this filter. No sequences affiliated with *Nitrosococcus* were detected. Thaumarchaeota (AOA; 2 ASVs) were detected at very low levels in sequence data (0.09% of total reads), though one or more ASVs were present in half of the DWTPs. Similar relative abundance levels were detected by qPCR, with an average absolute cell number of 1.1×10^6 /g of filter material (Figure 1D). With respect to other NOB, 6 ASVs affiliated with *Nitrotoga* spp. were observed in DWTP-2, 8, 9, and 12 at low abundance (<0.6%). No sequences were affiliated with *Nitrobacter* despite their detection by qPCR at around 1% of most communities (Figure 1D). Within the phylum Planctomycetes, several ASVs affiliated with anammox (order *Ca. Brocadiales*), but were detected at low levels (<0.06%) at only a few DWTPs (DWTP-8, 11, 12).

Comammox Nitrospira are the dominant nitrifiers

Sequencing and taxonomic assignment of *nxB* genes of *Nitrospira* spp. reveals that the filters are dominated by Lineage 2 *Nitrospira* (Figure 2A; 131 ASVs). Lineage 1 (8 ASVs) were present in DWTPs 2,5 and 8, while Lineage 4 (3 ASVs) was observed in DWTP-8 and 10. Eight unassigned ASVs were present at low abundance in DWTP-1 and 12. Though the high abundance of *Nitrospira* spp. in groundwater-fed RSFs was expected in accordance with previous work (Albers *et al.*, 2015; Gülay *et al.*, 2016), nothing is known about the diversity and abundance of comammox *Nitrospira* in the filters. To gain an idea of the relative proportion of comammox in the *Nitrospira* community, and to assess the accuracy of the newly designed *amoA* *Nitrospira* primers, we examined the taxonomic affiliation of 16S rRNA and *nxB* amplicon sequences from the filters (Table S1 & S2). The classification of comammox based on 16S rRNA and *nxB* is not robust due to high similarity of these genes in comammox and non-comammox *Nitrospira*, and has been applied here only as it was the sole means we could conceive of to evaluate *amoA* primer performance in real samples. Using strict similarity cutoffs with comammox reference sequences as criteria to consider our ASVs as comammox, (100% for 16S rRNA and 98% for *nxB*), comammox comprise a minimum of 78% (16S rRNA) and 28% (*nxB*) of the *Nitrospira* abundance in all filters (Table S1 & S2). The *nxB*-based estimation is likely very conservative, as almost one third (40 of 150 ASVs) of the *nxB* amplicon sequences did not affiliate closely (<91% identity) to any reference sequences, and the number of reference *nxB* sequences available for comammox is low. However 16S rRNA gene based estimates are in accordance with the qPCR data from the newly described *Nitrospira amoA* primers which, when normalized by total cell numbers of *Nitrospira* (from 16S rRNA and *nxB* gene qPCR), indicate that comammox make up between 28-100% of *Nitrospira* in the filters (Figure 1AB). Sequencing of *amoA* amplicons indicates that clade B comammox *Nitrospira* are more abundant than clade A, making up

around 75% of comammox abundance in most filters (Figure 2B). Despite this, there are both clade A and B ASVs that have high abundance in the majority of the filters, though in general, clade A appear to be less widespread across the filters (Figure 3). Phylogenetic
185 analysis of *amoA* sequences revealed a novel group of comammox within clade A that was present in a few filters for which no reference or clone sequences exist. This group's *amoA* is more closely related to canonical AOB than other known clade A comammox (Figure 4).

Ratios of nitrifying guilds were examined using qPCR data to identify dependencies between nitrifying groups. Linear regression showed a weak relationship between NOB and
190 AOP cell numbers, ($\text{NOB} = 2.5 \times \text{AOP}$, $R^2 \text{ adj} = 0.16$, $p = 0.045$). Unexpectedly, a strong relationship was observed between comammox *Nitrospira* and NOB ($\text{NOB} = 0.74 \times \text{Comammox}$, $R^2 \text{ adj} = 0.70$, $p = 1.5 \times 10^{-6}$) (Figure S5).

Impacts of physicochemical and operational parameters on nitrifying community

In order to explore the factors associated with the structure of the nitrifying
195 communities of the filters and the success of comammox *Nitrospira* in these communities, we performed constrained ordination of the nitrifying communities, using 16S rRNA of nitrifiers and *amoA Nitrospira* sequence data together with physicochemical composition of the influent water and filter material and operational parameters (Table S3). Constrained ordination models were built using stepwise selection of variables. Temperature was the
200 sole significant variable ($p = 0.025$) explaining variation in the nitrifying communities, explaining 9.4% of the variation. Increasing temperature appeared to be associated with increasing diversity of *Nitrospira* spp. in the filters (Figure 5A). Little explanatory power was provided by the measured physicochemical parameters to explain variation between the nitrifying communities. However, in ordination modelling of the comammox *Nitrospira* guild,
205 the sulfate content of the water, surface ammonium loading rate (SLR), nitrogen content of

the filter material and NVOC, were significant variables in addition to temperature explaining 20.5%, 13.7%, 13.4%, 11.5% and 10.4% of the variability respectively (Figure 5B, Table 2).

The ammonium concentration in the water and calcium concentration in the filter material were also found to be significant explanatory variables ($p < 0.04$) but were not included in
210 the ordination models due to co-linearity with other variables (Figure S6). Pearson correlations and linear regression were also used to relate the absolute abundance of nitrifying groups with physicochemical and operational parameters. The strongest correlation observed was between copper concentration in the influent water and comammox *Nitrospira* and *Nitrospira* cell numbers estimated from 16S rRNA qPCR ($R^2 = 0.63$
215 for both) (Figure S7, Table S4).

Discussion

Based on kinetic theory of optimal pathway length, Costa *et al.* (2006) predicted the existence of complete nitrifiers ten years before their discovery. They predicted that these
220 organisms would thrive in surface attached communities subject to low ammonium loading where slow growing, high yield organisms could flourish. Groundwater-fed RSFs represent just such an environment, and the abundance of comammox *Nitrospira* there confirms that this environment provides an ideal ecological niche for them. Similarly, comammox *Nitrospira* have been detected in relatively high abundance in other drinking water filters,
225 drinking water distribution systems and freshwater aquaculture filters (Pinto *et al.*, 2015; Bartelme *et al.*, 2017; Wang *et al.*, 2017). Developing an understanding of these organisms and their relevance to global nitrogen cycling requires investigations into their prevalence and contributions to nitrification in diverse environments. In order to facilitate these efforts,

we developed a method to detect and quantify diverse comammox *Nitrospira* of clades A
230 and B, and applied this method to a selection of RSFs.

These *amoA* primers cover the complete range of known comammox *Nitrospira*
diversity, with amplicons clustering with known reference sequences as well as presenting
newly identified sequences that cluster within both clades A and B (Figure 4). Complete
coverage of all sequences in a database of 48 *Nitrospira amoA* gene sequences when
235 allowing for 2 mismatches suggests excellent coverage of *Nitrospira amoA* with these
primers (Figure S4). One branch of *amoA* sequences within clade A did not cluster with any
reference sequences, and falls between *Ca. N. nitrosa* and Betaproteobacterial ammonia
oxidizers (Figure 4). This cluster was present in low abundance and in few filters, but is not
believed to be chimeric, as it clusters closely with other clade A sequences and these ASVs
240 were observed in several samples (Figure 3). It is difficult to assess whether the newly
designed *Nitrospira amoA* primers over- or underestimate comammox *Nitrospira* abundance,
though comammox cell numbers estimated by the assay tended to be lower or equivalent to
those estimated from 16S rRNA and *nxB Nitrospira* genes, making these estimates appear
reasonable. Comammox genomes assembled to date contain either one or two *amoA* genes,
245 however there could be a greater degree of variation in *amoA* copy number, which would
alter the estimation of comammox cell number based on *amoA* in the filter communities
(Figure 1). It can also not yet be evaluated whether the assay is biased towards a specific
clade, as the true relative abundance of clades A and B in the sampled RSFs is not clear.
Enrichment and genome assembly of additional comammox *Nitrospira* and further
250 investigation into the breadth of comammox diversity will provide additional insights into
the coverage of the qPCR assay. Several other primers for the detection of comammox have
been developed. One approach uses a two-step method with primers that have broad

coverage of copper-containing membrane-bound monooxygenases (CuMMOs) combined with a newly designed specific primer for comammox (Wang *et al.*, 2017). While this method has a broad coverage for CuMMOs, it does not have good coverage of comammox *amoA*, and due to its two-step nature, is not amenable to quantification. Another set of primers has high coverage and specificity, but requires separate primer sets for clade A and clade B (Pjevac *et al.*, 2017). A third group has combined a previously designed *pmoA* primer with a newly designed comammox specific primer. The coverage of this primer was not investigated, and the study amplified only a single comammox sequence. Whether this was due to low coverage of the primers, or low diversity of comammox in the examined samples is not clear (Bartelme *et al.*, 2017). Thus, we believe that our primers fill an important gap, providing broad coverage amplification and quantification of both comammox clades with a single primer set.

The nitrifying communities of all filters were dominated by the phylum Nitrospirae. Members of the Nitrospirae were mostly from the genus *Nitrospira* (112/123 ASVs), and these are believed to catalyse either complete nitrification or nitrite oxidation (Figure 1). Analysis of *nxB* gene sequences from the filter communities indicate that the majority of the *Nitrospira* present are from Lineage 2 (Figure 2A), and classification of *nxB* sequences (at $\geq 98\%$ nucleotide identity to reference sequences) indicates that a minimum of 28% of the *Nitrospira* affiliate with comammox *Nitrospira*, while qPCR results from 16S rRNA and *amoA* from *Nitrospira* suggests that this value is much higher (40-100%) (Tables S2 & S3). Nitrosomonadaceae, dominated by *Nitrosomonas* spp. were present in all filters but generally in low abundance (<2.5%). A very small proportion (<0.06%) of AOA were present in some filters. *Nitrobacter* spp., while not detected in sequence analysis, were detected by qPCR, which suggested that their abundance was similar, or slightly lower than that of AOB

(Figure 1D). Inter-filter variation was observed in the nitrifying communities. In certain filters, notably DWTP-9 and 12, lower *Nitrospira* abundance was associated with higher Nitrosomonadaceae. In addition, DWTP-12 also had a larger proportion of *Nitrotoga* spp. than any other filter. These observations suggest that functional redundancy exists in the filters, and that a functional nitrifying community was present in all filters. While the community structures were similar to those in previously examined filters, our study found a somewhat higher fraction of *Nitrospira* than one previous study (Albers *et al.*, 2015) but is in line with *Nitrospira* abundances in a second study (Gülay *et al.*, 2016). Nitrosomonadaceae estimates are consistent with results from both studies and neither detected *Nitrobacter* spp. in amplicon data. Interestingly, the same primers and sequencing platforms (454 pyrosequencing) were used in these two studies, ruling out possibilities of primer bias to explain diverging *Nitrospira* estimates. Sequencing depth was greater (by at least 2 fold) in the second study. However these results likely suggest that different filters contain varying abundances of *Nitrospira* spp. Several different filter types were investigated in the present study including pre-filters (PF) (DWTP-7-PF and DWTP-11), single filters (SF) (DWTP-1-3 and 12) and after-filters (AF) (DWTP-4-10). In groundwater-fed RSFs, PFs are generally presumed to be the major site of abiotic reactions such as iron oxidation and precipitation, and generally contain a larger grain size, while AFs are considered the main site of nitrification, and contain smaller grain sizes (Tatari *et al.*, 2017). The low abundance of nitrifiers at DWTP-11 agrees with a dominant abiotic role for PFs, but the large nitrifying community at DWTP-7-PF does not support this. However, DWTP-7 is a special case where the PF is particularly deep (2.3 m) and is the main site of ammonium removal (Søborg *et al.*, 2015).

Based on experimental observations and energy conservation ratios, the ratio of AOP to NOB in a nitrifying reactor is around 2:1 (Winkler *et al.*, 2012). Until their discovery, the

presence of comammox *Nitrospira* would result in an apparent overabundance of NOB, which would typically be attributed to a nitrification-denitrification loop, or a so-called ping-pong effect (Winkler *et al.*, 2012). Based on the stoichiometry of nitrification and the presumed metabolisms of comammox, AOP and NOB, in an environment containing all three

305 guilds it is expected that that there would not be a fixed ratio of comammox to AOP and NOB, while the ratios of AOP and NOB would remain as expected. We calculated ratios of these three groups across the rapid sand filters based on qPCR data where all groups were quantifiable and conducted linear regression of the relative abundance of nitrifying guilds (Figure S5). A weak linear relationship was observed between AOP and NOB ($\text{NOB} = 2.5 \times$

310 AOP , $R^2 \text{ adj} = 0.16$, $p = 0.045$), and NOB are much more abundant than expected if all energy comes from nitrite produced by AOP. In addition, a strong relationship exists between comammox *Nitrospira* and NOB ($\text{NOB} = 0.74 \times \text{Comammox}$, $R^2 \text{ adj} = 0.70$, $p = 1.49 \times 10^{-6}$, Figure S5). These observations do not follow expectations based on the assumptions that comammox *Nitrospira* oxidize ammonium completely to nitrate and that canonical AOB and

315 NOB then share the ammonium not oxidized by comammox *Nitrospira*. Several factors could contribute to these observations. The first is that the newly designed primers underestimate the number of comammox, and that a portion of the organisms considered to be NOB are actually comammox. However, this would require a very large underestimation of comammox by these primers to return the AOB:NOB ratio close to 2:1. The presence of a

320 nitrification-denitrification loop is certainly a possibility in these environments – a number of the abundant organisms in these filters such as members of the Bacteroidetes, and Chloroflexi are potentially capable of nitrate reduction, which could provide additional nitrite for NOB. Some NOB may also have diverse metabolic capabilities, and are not solely oxidizing nitrite, as has been found in some *Nitrospira* (Koch *et al.*, 2014, 2015), though the

325 potential alternate substrates that would provide sufficient energy in these filters are
unclear. A fourth possibility, that explains the strong relationship between NOB and
comammox is that comammox *Nitrospira* support the growth of NOB by providing them with
a fraction of the nitrite they produce. This would be consistent with the observed transient
accumulation of up to 30% of the nitrite in enriched cultures of *Ca. Nitrospira inopinata*
330 during ammonia oxidation (Daims *et al.*, 2015). At micromolar ammonium concentrations,
which are reflective of the concentrations in the top of most of the filters, nitrite
accumulation was lower, but was still substantial (15-20%). This suggests that comammox
Nitrospira may feed canonical NOB, thereby promoting their growth. There is no clear
reason for a metabolic dependency of, or benefits of cooperation for, comammox with NOB,
335 and it is possible that this phenomenon is due solely to the diffusion of nitrite out of the
periplasmic space (where it is generated) during nitrification. In *Ca. Nitrospira inopinata*, the
only comammox organism for which data is available, the affinity for nitrite is three orders of
magnitude lower than that of ammonia ($K_m = 449.2 \mu\text{M}$) (Kits *et al.*, 2017), and is
substantially lower compared with characterized canonical nitrite-oxidizing *Nitrospira* ($K_m =$
340 $9-27 \mu\text{M}$) (Nowka *et al.*, 2015). This relatively lower affinity for nitrite than ammonia in
comammox suggests that nitrite could accumulate in the periplasm of comammox during
nitrification, and subsequently diffuse outside of the cell where it could be used by nearby
NOB. *In-situ* visualization of nitrifying biofilms would indicate if NOB are found in close
association with comammox *Nitrospira* in order to take advantage of nitrite diffusion. This
345 would provide an additional explanation for the frequent co-existence of different *Nitrospira*
spp. in diverse environments (Gruber-Dorninger *et al.*, 2015; Gülay *et al.*, 2016).

Due to the high similarities in physicochemical parameters and microbial
communities between filters, correspondence analysis did not provide strong links between

physicochemical parameters and the total nitrifying community or comammox *Nitrospira* guild structure (Figure 5, Table 2). However, our analysis does suggest that temperature influences the structure of the nitrifying community (Figure 5A). Temperatures in groundwater-fed RSFs tend to be low, and are fairly consistent year-round due to low fluctuations in subsurface temperature. The temperature in the studied filters ranged from 8.6-13.2°C, which creates a selective pressure for organisms that are active and growing at low temperature. Increasing temperature appears to have a positive effect on *Nitrospira* spp., though we could not distinguish between comammox and nitrite-oxidizing taxa (Figure 5A). In addition to temperature, the sulfate content of the influent water, surface ammonium loading rate, nitrogen content of the filter material and NVOC, play significant roles in shaping the comammox *Nitrospira* guild in RSFs. However, there was no clear clade-specific separation within comammox *Nitrospira* based on these variables. As all source waters are freshwater, sulfate was extremely low in all environments (<25 mg/L), so it was interesting to find that it was an important explanatory variable in shaping the comammox guild, even at such low concentrations. Notably, while comammox *Nitrospira* have been detected in brackish lakes and marine sediments (van Kessel *et al.*, 2015; Pjevac *et al.*, 2017), they have mainly been described in freshwater environments (Pinto *et al.*, 2015; Palomo *et al.*, 2016; Y. Wang *et al.*, 2017) and have not been detected in marine environments (Daims *et al.*, 2015). Thus, it is possible that sensitivity to sulfate varies across the comammox *Nitrospira*, even at the low concentrations observed in freshwater environments.

Investigations of sulfur metabolism in comammox genomes may shed light on their sensitivity to sulfate. Ammonium is presumed to be the primary biological substrate in the majority of the filters examined here, so the impact of the surface ammonium loading rate on community structure was not surprising. Low ammonium flux is one of the factors

predicted to select for comammox based on the kinetic theory of optimal pathway length (Costa *et al.*, 2006), and our observations suggest that this factor is important even within
375 the comammox guild, suggesting that there may be differential response to ammonium loadings within comammox, though no clade-specific differentiation was observed. Ammonium concentration, which was also observed to be significant in explaining the structure of the comammox community but was not included in the associated model, is strongly correlated with the surface ammonium loading rate (Figure S6). The majority of
380 nitrogen in the filter material is found in amino form (around 70% in all filters). High resolution scanning of the carbon (C1s), nitrogen (N1s), and oxygen (O1s) regions points towards the presence of amino acids and/or protein at all measured depths of the mineral coatings (Figure S8). It is possible that proteins or amino acids are produced during biofilm formation by comammox *Nitrospira* spp., explaining the relationship between comammox
385 and filter material nitrogen. Alternately, they may be produced by other community members, but are responded to by certain comammox *Nitrospira*. The excretion of proteins and amino acids during biofilm formation has been observed previously in diverse taxa (Flemming and Wingender, 2010). NVOC explained significant variation only in the comammox community. The presence of mixotrophy within the genus *Nitrospira* (Watson *et al.*, 1986; Daims *et al.*, 2001; Spieck *et al.*, 2006) suggests that some comammox *Nitrospira*,
390 like their nitrite-oxidizing counterparts, could be capable of mixotrophic growth, resulting in selection of specific strains based on organic carbon availability. Investigation of additional filters with more variable microbial communities and physicochemical parameters may provide further information to support or refute these observations and provide stronger
395 links between environmental parameters and comammox *Nitrospira* ecology. Physiological analysis of enriched or pure strains, as well as genomic analysis of comammox *Nitrospira* will

provide further insight into the role of these features in comammox *Nitrospira*. Linear regression between the abundance of nitrifying clades with physicochemical parameters resulted in a significant linear relationship between comammox abundance and the copper content of the water, though copper was not identified as an important factor in ordination analysis. As copper is an essential co-factor in the ammonia monooxygenase enzyme, it is not surprising that it could limit the growth of ammonia oxidizers. Recovery of nitrification activity in poorly functioning RSFs after copper dosing has been previously observed (Wagner *et al.*, 2016).

Despite having been only recently described, the wide environmental distribution of comammox *Nitrospira* as detected in metagenomic databases suggests that comammox *Nitrospira* likely make a significant contribution to global nitrogen cycling (Daims *et al.*, 2015). The developed primers for the *amoA* gene of comammox *Nitrospira* exhibit good coverage of both clades A and B, and enabled the quantification and characterization of comammox *Nitrospira* across 12 DWTPs. Comammox *Nitrospira* are by far the most abundant nitrifiers in all groundwater-fed RSFs examined. Though the specific activity of comammox *Nitrospira* in the filters is not known, they likely play a major role in nitrification in the filters as the abundance of other ammonia and nitrite oxidizers was comparatively low. Furthermore, a relationship between the abundance of comammox and NOB suggests that comammox *Nitrospira* oxidize a greater proportion of ammonia entering the filters than AOB, and may directly support the growth of NOB. Future work involving isotope labelling, and enrichment and isolation of comammox *Nitrospira* will provide clarity as to their specific nitrification activity, and their contribution to global nitrogen cycling.

420 **Experimental Procedures**

Sampling

The twelve drinking water treatment plants (DWTPs) are geographically distributed across Denmark (Figure S9). Physicochemical characteristics of influent, effluent water and sand were obtained through direct measurement or from the JUPITER database

425 (www.geus.dk/DK/data-maps/jupiter/Sider/default.aspx) (Table S3). Filter material (15 mL) was collected from 2 locations at the top of the filters using a 1% hypochlorite-wiped stainless steel grab sampler. Single sand samples were collected from DWTP-7, 9, 10, 11, 12. Filter material was immediately placed into cryotubes, immersed in liquid nitrogen, and stored at -80 °C for further analysis. Liquid grab samples were collected from i) the pre-filter
430 influent from the aeration steps ii) the after-filter influent from the after-filter cascade and iii) the effluent tap at the end of the after-filter. Liquid samples were collected in 1 L glass bottles and were filter sterilized (0.2 µm) into polypropylene bottles for specific chemical analyses. Samples for ICP-MS/ OES were immediately acidified (65% nitric acid, Merck, Suprapur). Liquid samples were transported on ice and stored at -20 °C until further analysis.

435 *Chemical analysis of influent and effluent water and filter material*

Ammonium was measured using a standard colorimetric salicylate and hypochlorite method (Bower and Holm-Hansen, 1980). Nitrite was analysed using a standard method adapted from Grasshoff *et al.* (1983). Nitrate and sulfate were measured by ion chromatography according to AWWA-WEF method 4110 (Eaton *et al.*, 1998). Metal content
440 in water was determined by ICP-MS (Fe, Mn, Cu, Zn, P, Mg, Co, and Ni; 7700x, Agilent Technologies), or ICP-OES (Ca; Varian, Vista-MPX CCD Simultaneous ICP-OES). Dissolved oxygen and pH were measured with a handheld meter (WTW, Multi 3430, with FDO® 925 and SenTix® 940 probes). NVOC analysis was performed using a wet chemical TOC-analyser TOC-V WP (Shimadzu, Kyoto, Japan). Dried sand samples (60°C, 16 hours) were used for

445 mineral coating characterization and elemental composition by X-ray photoelectron spectroscopy (XPS, K-Alpha, Thermo Scientific, USA). Triplicate sand grains were analysed under vacuum with a monochromated Al-K α X-ray source and an X-ray spot size of 400 μ m.

Nitrospira amoA qPCR primer and assay design

To design primers for *amoA Nitrospira*, DNA sequences for all high quality *Nitrospira*
450 *amoA* genes available at that time (n = 8) were collected and aligned using T-coffee (Notredame *et al.*, 2000). Alignments were manually curated, and regions of homology between sequences were further examined as candidate regions for primer design. The criteria used to evaluate and develop primer sets were that the primer region was between 16 and 24 nucleotides long with a GC content close to 50%, contained minimal secondary
455 structure and a minimal number of non-homologous bases between sequences in the alignment including no non-matching bases within 4 bases of the 3' end. In cases of non-matching nucleotides, the possibility of improving primer coverage by using degenerate bases was examined, maximizing the use of degenerate bases that would result in the highest likelihood of amplification of all sequences. Candidate primers that would result in
460 amplicons of 100 to 350 bp were examined for compatibility (minimal T_m difference and tendency for dimer formation). To evaluate specificity, primer sets were searched against the NCBI non-redundant nucleotide database. No non-specific binding results were detected for the final primer set. Primers Ntsp-amoA 162F (GGATTCTGGNTSGATTGGA) and Ntsp-amoA 359R (WAGTTNGACCACCASTACCA) were synthesized and tested *in vitro* against
465 strains of *E. coli* (TOP10, Invitrogen), *Pseudomonas putida* KT2440, *Clostridium thermocellum*, *Methanoculleus bourgensis* and *Meliobacter roseus* as well as a number of bioreactor and environmental samples (nitritation anammox reactors, DWTPs, denitrifying reactors). The annealing temperature was optimized using gradient PCR (46°C-50°C) with negative and

positive control samples. With the final amplification protocol, single bands of the expected
470 size were observed only in positive samples. A clone library was made by pooling amplicons
from 3 DWTPs (DWTP-8, 12 and Glostrup waterworks, Zealand), which were used in a TOPO
TA cloning reaction according to manufacturer's instructions (Invitrogen). Thirty positive
clones were picked and were amplified with *amoA Nitrospira* and M13 primer sets, all
resulting in the expected band size. From these clones, 21 plasmids were purified using the
475 Qiagen Plasmid MiniPrep kit (QIAGEN, Germany) and sequenced by MacroGen Inc. on an
ABI3730XL sequencer from the M13R primer (Amsterdam, NL). Plasmids were subsequently
used for standard curves for absolute quantification by qPCR. The standard curve is linear
between 10^{-10} to 10^9 copies/reaction and the efficiency of the assay was around 105%. PCR
reactions were carried out in a volume of 25 μ L consisting of 5X Phusion HF buffer (Thermo
480 Fisher), 0.5 μ L of 10 mM dNTPs (Sigma), 500 nM of each primer, 0.25 μ L of Phusion DNA
polymerase (ThermoFisher) and 1 μ L of DNA template in DNA/RNA free water with the
following thermocycling protocol: 95°C for 10 min, followed by 25 cycles of 95°C for 45 s,
48°C for 30 s, 72°C for 45 s, and final extension at 72 °C for 7 min. Gene fragments of
Nitrospira amoA from clones have been deposited in Genbank (MF073209-MF073227). A
485 database of available *amoA Nitrospira* sequences was constructed using published reference
sequences and mining of the IMG database. Phylogenetic trees were constructed using
known *Nitrospira* and *Nitrosomonas amoA* sequences to distinguish *Nitrospira amoA*
sequences (Figure S4). Primer coverage was assessed *in silico* using R package BioStrings.

DNA extraction and qPCR

490 DNA was extracted from 0.5 g of sand material using the MP FastDNA Spin Kit (MP
Biomedicals LLC, Solon, USA) as described (Palomo *et al.*, 2016). DNA concentration and
quality was measured by NanoDrop (NanoDrop Technologies, Wilmington, USA). Gene copy

numbers of 16S rRNA from total Bacteria, AOB, and *Nitrospira*, *amoA* from Archaea, and *Nitrospira*, and *nxB* from *Nitrospira* and *Nitrobacter* were estimated by qPCR (Table 1).

495 Reactions were carried out on a Chromo4 thermocycler (Bio-Rad) in a total volume of 25 μ L containing 10 ng of DNA, 12.5 μ L of 2x iQ SYBR Green Supermix (Bio-Rad), and 500 nM of each primer in DNA/RNA free water. Gene copies were quantified by comparison to a standard curve constructed with plasmid dilutions at known copy numbers. Copy numbers of 16S rRNA from Bacteria were converted to cell numbers based on the average 16S rRNA
500 gene copy number per organism in each sample by comparing 16S rRNA amplicon data to the rrnDB database (Stoddard *et al.*, 2015) and calculating the average copy number per sample (R script available at <https://github.com/ardagulay>). Other qPCR data were corrected using the following assumptions: *Nitrobacter* spp. contain 2 copies of *nxB*, while *Nitrospira* spp. contain one copy. A single copy of *amoA* per genome is present in *Nitrospira* and AOA,
505 and similarly a single copy of 16S rRNA is present in *Nitrospira* and AOB.

Amplicon sequencing and analysis

DNA was amplified using primers for 16S rRNA total Bacteria, *nxB Nitrospira* and *amoA Nitrospira* (Table 1). Library preparation and Illumina MiSeq sequencing was carried out by DMAC (DTU Multi-Assay Core Facility, Kgs Lyngby, DK). Illumina data was analysed in
510 USEARCH to examine quality in order to optimize trimming procedures (Edgar, 2010). Quality control, trimming, merging of paired ends, and error correction were performed in DADA2, which outputs the abundance of error-corrected amplicon sequence variants (ASVs) rather than clustering as in an OTU-based approach (Callahan *et al.*, 2016). Comparison of ASVs with the SILVA SSU database v123 was used for taxonomic assignment of 16S rRNA
515 genes, while custom databases were used for taxonomic assignment of *nxB* and *amoA* of *Nitrospira*. Further analysis was carried out in R package phyloseq (2012; McMurdie *et al.*,

2013). Ratios of nitrifiers were calculated using qPCR data – NOB abundance was inferred by subtracting *Nitrospira amoA* copy numbers from *Nitrospira nxrB* copy numbers and adding *Nitrobacter* cell numbers. AOP were inferred by adding AOB 16S rRNA copy numbers and
520 AOA *amoA* copy numbers. Amplicon data are available under PRJNA399693 (Table S5).

Statistical analysis

Constrained correspondence analysis (CCA) with variation partitioning, linear regression and correlation analyses were performed in R package *vegan* (Oksanen *et al.*, 2017). CCA was performed using ASV abundances of 16S rRNA of the nitrifying fraction of
525 the community, (family Nitrosomonadaceae, *Can. Nitrotoga*, *Can. Nitrosoarchaeum*, and genus *Nitrospira*) or *amoA* amplicons together with filter operational parameters and physicochemical data of influent water and filter material. Libraries were normalized to relative abundance. In the analysis of the comammox *Nitrospira* guild structure, all *amoA* ASVs were used in the analysis, but libraries were rarefied to even depth (30799 reads). CCA
530 was performed to determine the importance of each variable, and a constrained ordination model was built using stepwise selection of variables with *ordistep*. Variable significance was determined by permutation tests (200 permutations) with *anova.cca* in *vegan*.

Acknowledgements

535 The authors would like to acknowledge the support of George Kwarteng Amoako (HOFOR), Stig Eskildsen (Forsyning Ballerup), Steffen Schulz and Jørgen Bendtsen (Nordvand), Torben Snehøj Weng (Fors A/S), and Jens Kjølby Larsen (Forsyningen) for their support in sample and data collection, as well as Carson Odell Lee, Florian B. Wagner and Hans-Jørgen Albrechtsen for providing samples and data for some of the waterworks, and Marlene
540 Danner Dalgaard and DMAC for sequencing support. This research was supported by

MERMAID: An initial training network funded by the People Programme- Marie Skłodowska-Curie Actions of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement no. 1607492 and by a research grant (13391, Expa-N) from VILLUM FONDEN.

References

- Albers, C.N., Ellegaard-Jensen, L., Harder, C.B., Rosendahl, S., Knudsen, B.E., Ekelund, F., and
550 Amand, J. (2015) Groundwater chemistry determines the prokaryotic community
structure of waterworks sand filters. *Environ. Sci. Technol.* **49**: 839–846.
- 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.*
555 **8**: 101.
- Bower, C.E. and Holm-Hansen, T. (1980) A Salicylate–Hypochlorite Method for Determining
Ammonia in Seawater. *Can. J. Fish. Aquat. Sci.* **37**: 794–798.
- Callahan, B.J., Mcmurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J., and Holmes, S.P. (2016)
DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods*
560 **13**: 581.
- Costa, E., Pérez, J., and Kreft, J.-U. (2006) Why is metabolic labour divided in nitrification?
Trends Microbiol. **14**: 213–219.
- 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.
- 565 Daims, H., Nielsen, J.L., Nielsen, P.H., Schleifer, K.H., and Wagner, M. (2001) In Situ
Characterization of Nitrospira-Like Nitrite-Oxidizing Bacteria Active in Wastewater
Treatment Plants. *Appl. Environ. Microbiol.* **67**: 5273–5284.
- Eaton, A., Clesceri, L., Greenberg, A., Franson, M., American Public Health Association,
American Water Works Association, and Water Environment Federation (1998)

- 570 Standard methods for the examination of water and wastewater 20th ed. American
Public Health Association, Washington D.C.
- Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST.
Bioinformatics **26**: 2460–1.
- Ferris, M.J. and Muyzer, G. (1996) Denaturing Gradient Gel Electrophoresis Profiles of 16S
575 rRNA-Defined Populations Inhabiting a Hot Spring Microbial Mat Community. **62**: 340–
346.
- Flemming, H. and Wingender, J. (2010) The biofilm matrix. *Nat. Rev. Microbiol.* **8**: 623–33.
- Graham, D.W., Knapp, C.W., Van Vleck, E.S., Bloor, K., Lane, T.B., and Graham, C.E. (2007)
Experimental demonstration of chaotic instability in biological nitrification. *ISME J.* **1**:
580 385–393.
- Grasshoff, K., Ehrhardt, M., and Kremling, K. eds. (1983) *Methods of Seawater Analysis* 2nd
ed. Verlag Chemie, Basel, CH.
- Gruber-Dorninger, C., Pester, M., Kitzinger, K., Savio, D.F., Loy, A., Rattei, T., et al. (2015)
Functionally relevant diversity of closely related *Nitrospira* in activated sludge. *ISME J.* **9**:
585 643–655.
- Gülay, A., Musovic, S., Albrechtsen, H.-J., Al-Soud, W.A., Sørensen, S.J., and Smets, B.F. (2016)
Ecological patterns, diversity and core taxa of microbial communities in groundwater-
fed rapid gravity filters. *ISME J.* **10**: 2209–2222.
- Hermansson, A. and Lindgren, P.E. (2001) Quantification of ammonia-oxidizing bacteria in
590 arable soil by real-time PCR. *Appl. Environ. Microbiol.* **67**: 972–6.

- van Kessel, M.A.H.J., Speth, D.R., Albertsen, M., Nielsen, P.H., Op den Camp, H.J.M., Kartal, B., et al. (2015) Complete nitrification by a single microorganism. *Nature* **528**: 555–559.
- Kits, K.D., Sedlacek, C.J., 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.
- 595 Koch, H., Galushko, A., Albertsen, M., Schintlmeister, A., Gruber-Dorninger, C., Lucker, S., et al. (2014) Growth of nitrite-oxidizing bacteria by aerobic hydrogen oxidation. *Science* (80-.). **345**: 1052–1054.
- Koch, H., Lücker, S., Albertsen, M., Kitzinger, K., Herbold, C., Spieck, E., et al. (2015) Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus
600 *Nitrospira*. *Proc. Natl. Acad. Sci.* **112**: 11371–11376.
- Könneke, M., Bernhard, A.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B., and Stahl, D. a (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543–6.
- Kowalchuk, G.A., Stephen, J.R., De Boer, W., Prosser, J.I., Embley, T.M., and Woldendorp, J.W.
605 (1997) Analysis of Ammonia-Oxidizing Bacteria of the γ Subdivision of the Class Proteobacteria in Coastal Sand Dunes by Denaturing Gradient Gel Electrophoresis and Sequencing of PCR-Amplified 16S Ribosomal DNA Fragments. *Appl. Environ. Microbiol.* **63**: 1489–1497.
- Lane, D.J. (1991) 16S/23S rRNA sequencing. In, Stackebrandt, E. and Goodfellow, M. (eds),
610 *Nucleic acid techniques in bacterial systematics*. John Wiley and Sons, New York, NY, pp. 115–175.

- McMurdie, P.J., Holmes, S., Kindt, R., Legendre, P., and O'Hara, R. (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**: e61217.
- 615 Notredame, C., Higgins, D.G., and Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**: 205–217.
- Nowka, B., Off, S., Daims, H., and Spieck, E. (2015) Improved isolation strategies allowed the phenotypic differentiation of two *Nitrospira* strains from widespread phylogenetic lineages. *FEMS Microbiol. Ecol.* **91**: 1–11.
- 620 Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., et al. (2017) vegan: Community Ecology Package. R package version.
- Palomo, A., Jane Fowler, S., Gülay, 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.
- 625 Pester, M., Maixner, F., Berry, D., Rattei, T., Koch, H., Lüscher, 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.
- Pinto, A.J., Marcus, D.N., Ijaz, Z., Bautista-de los Santos, Q.M., Dick, G.J., and Raskin, L. (2015) Metagenomic Evidence for the Presence of Comammox *Nitrospira*-Like Bacteria in a
630 Drinking Water System. *mSphere* **1**: e00054-15.
- Pjevac, P., Schaubberger, C., Poghosyan, L., Herbold, C.W., van Kessel, M.A.H.J., Daebeler, A., et al. (2017) AmoA-targeted polymerase chain reaction primers for the specific detection and quantification of comammox *Nitrospira* in the environment. *bioRxiv*.

- Søborg, D.A., Breda, I.L., and Ramsay, L. (2015) Effect of oxygen deprivation on treatment
635 processes in a full-scale drinking water biofilter. *Water Sci. Technol. Water Supply* **15**:
825–833.
- Spieck, E., Hartwig, C., McCormack, I., Maixner, F., Wagner, M., Lipski, A., and Daims, H.
(2006) Selective enrichment and molecular characterization of a previously uncultured
Nitrospira-like bacterium from activated sludge. *Environ. Microbiol.* **8**: 405–415.
- 640 Stoddard, S.F., Smith, B.J., Hein, R., Roller, B.R.K., and Schmidt, T.M. (2015) rrnDB: improved
tools for interpreting rRNA gene abundance in bacteria and archaea and a new
foundation for future development. *Nucleic Acids Res.* **43**: D593–D598.
- Tatari, K., Musovic, S., Gülay, A., Dechesne, A., Albrechtsen, H.-J., and Smets, B.F. (2017)
Density and distribution of nitrifying guilds in rapid sand filters for drinking water
645 production: dominance of Nitrospira spp. *Water Research* **127**: 239–248.
- Team, R.C. (2012) R: A language and environment for statistical computing.
- Tourna, M., Freitag, T.E., Nicol, G.W., and Prosser, J.I. (2008) Growth, activity and
temperature responses of ammonia-oxidizing archaea and bacteria in soil microcosms.
Environ. Microbiol. **10**: 1357–1364.
- 650 Vanparrys, B., Spieck, E., Heylen, K., Wittebolle, L., Geets, J., Boon, N., and De Vos, P. (2007)
The phylogeny of the genus Nitrobacter based on comparative rep-PCR, 16S rRNA and
nitrite oxidoreductase gene sequence analysis. *Syst. Appl. Microbiol.* **30**: 297–308.
- Wagner, F.B., Nielsen, P.B., Boe-Hansen, R., and Albrechtsen, H.J. (2016) Copper deficiency
can limit nitrification in biological rapid sand filters for drinking water production.
655 *Water Res.* **95**: 280–288.

- Wang, J.-G., Xia, F., Zeleke, J., Zou, B., Rhee, S.-K., and Quan, Z.-X. (2017) An improved protocol with a highly degenerate primer targeting copper-containing membrane-bound monooxygenase genes for community analysis of methane- and ammonia-oxidizing bacteria. *FEMS Microbiol. Ecol.* **93**: fiw244.
- 660 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.
- Watson, S.W., Bock, E., Valois, F.W., Waterbury, J.B., and Schlosser, U. (1986) *Nitrospira marina* gen. nov. sp. nov.: a chemolithotrophic nitrite-oxidizing bacterium. *Arch. Microbiol.* **144**: 1–7.
- 665 Winkler, M.K.H., Bassin, J.P., Kleerebezem, R., Sorokin, D.Y., and Van Loosdrecht, M.C.M. (2012) Unravelling the reasons for disproportion in the ratio of AOB and NOB in aerobic granular sludge. *Appl. Microbiol. Biotechnol.* **94**: 1657–1666.
- Yu, Y., Lee, C., Kim, J., and Hwang, S. (2005) Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction.
- 670 *Biotechnol. Bioeng.* **89**: 670–9.

Figures

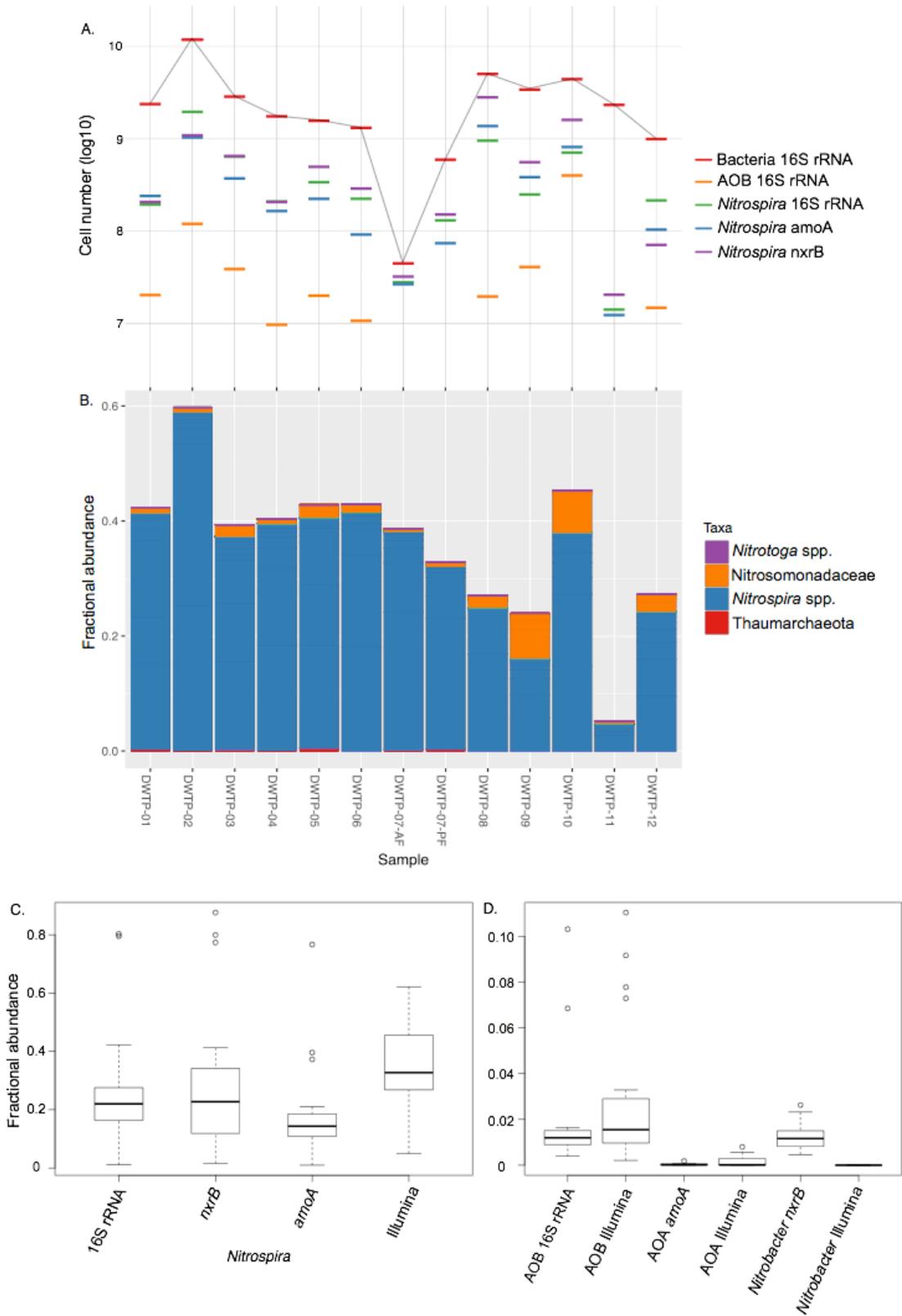
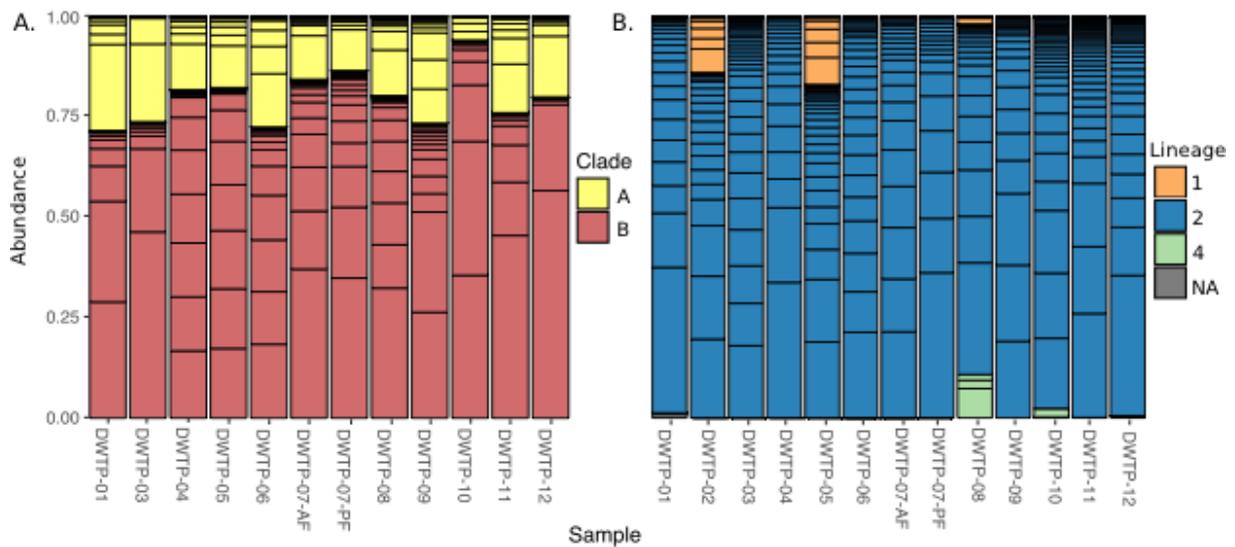


Figure 1. Characterization of nitrifying communities from 12 groundwater-fed RSFs based on a) qPCR and b) 16S rRNA gene amplicon sequencing. c) Comparison of *Nitrospira* relative abundance by qPCR and 16S rRNA amplicon sequencing. d) Comparison of AOB, AOA (*amoA*) and *Nitrobacter* (*nrxB*) relative abundance based on qPCR and 16S rRNA amplicon sequencing. In c) and d) qPCR cell numbers were normalized to relative abundance based on 16S rRNA Bacteria-derived cell numbers

680



685

Figure 2. Relative abundances of *Nitrospira* spp. ASVs by a) lineage, based on *nrxB* sequencing, b) comammox clade based on *amoA* sequencing

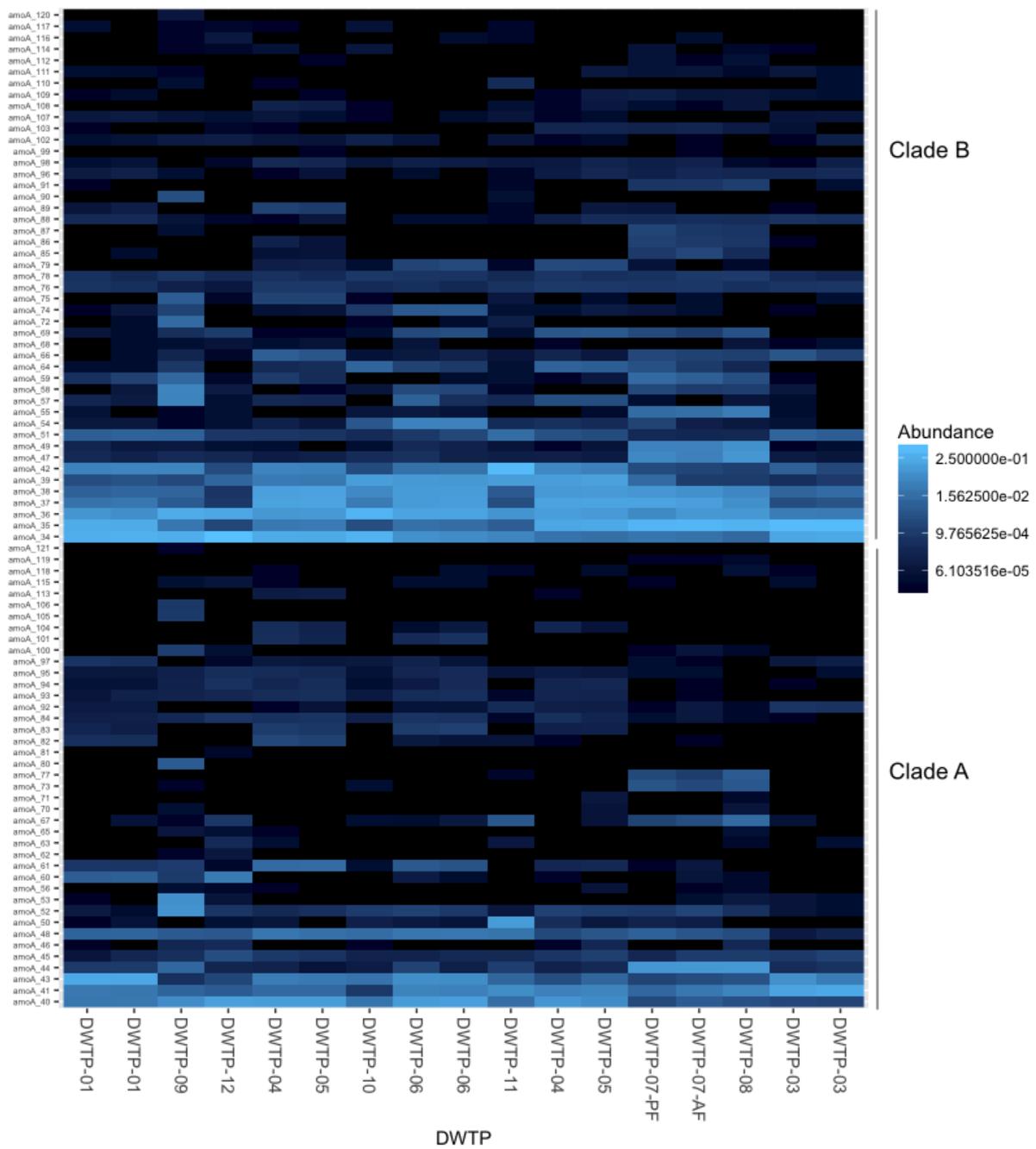
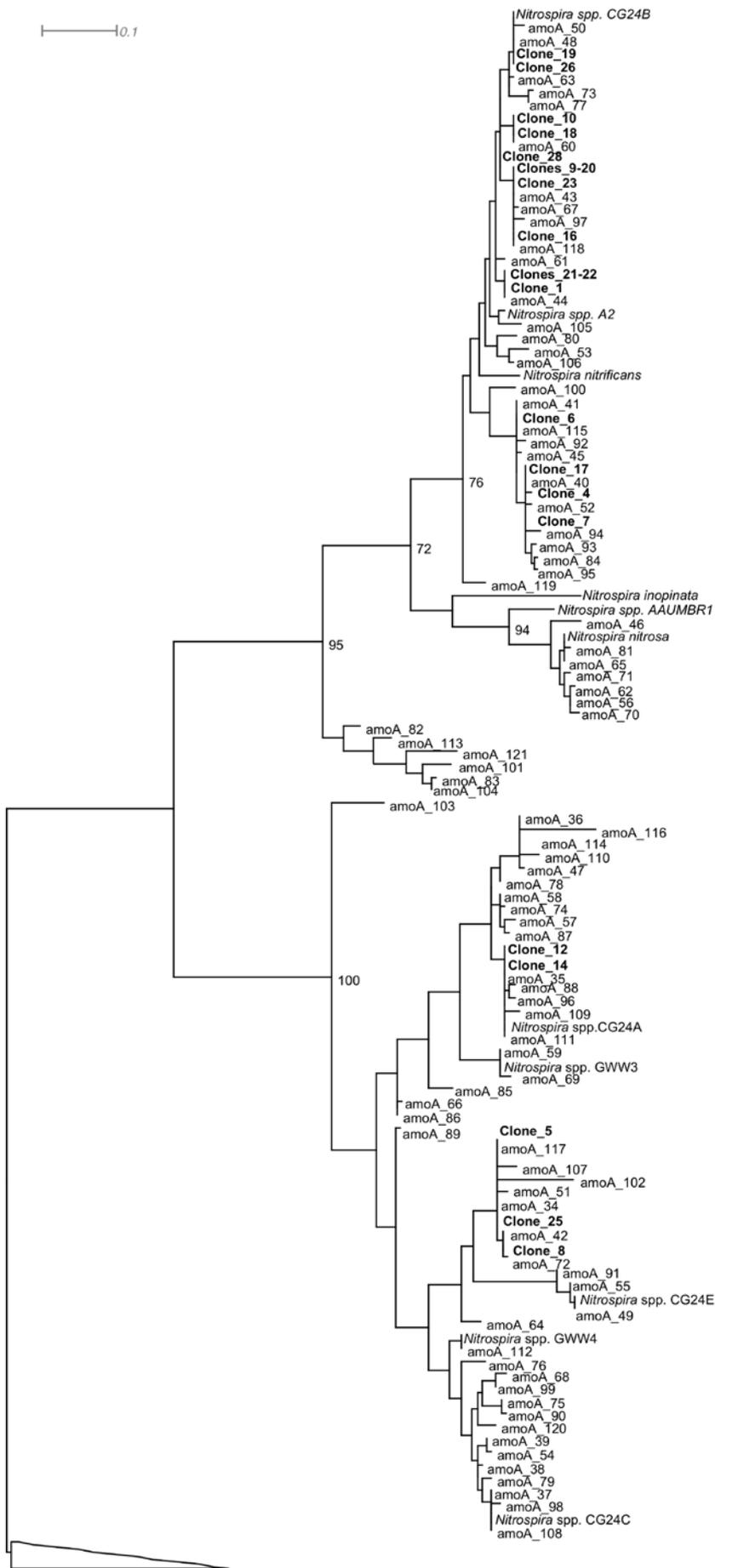


Figure 3. Heatmap showing the relative fractional abundance of *amoA* *Nitrospira* amplicon sequences across eleven waterworks.

0.1



Clade A

Clade B

695 **Figure 4. Maximum likelihood tree of *amoA Nitrospira* amplicons and reference nucleotide sequences and clone sequences from this study. Sequences were subject to frameshift correction and codons were aligned using DECIPHER. Phylogenetic analysis was performed in MEGA7. Bootstrap values greater than 60 are shown. The tree was rooted using *amoA* sequences from *Nitrosomonas europaea*, *Nitrosomonas eutropha* and *Nitrosomonas sp. Is79A3***

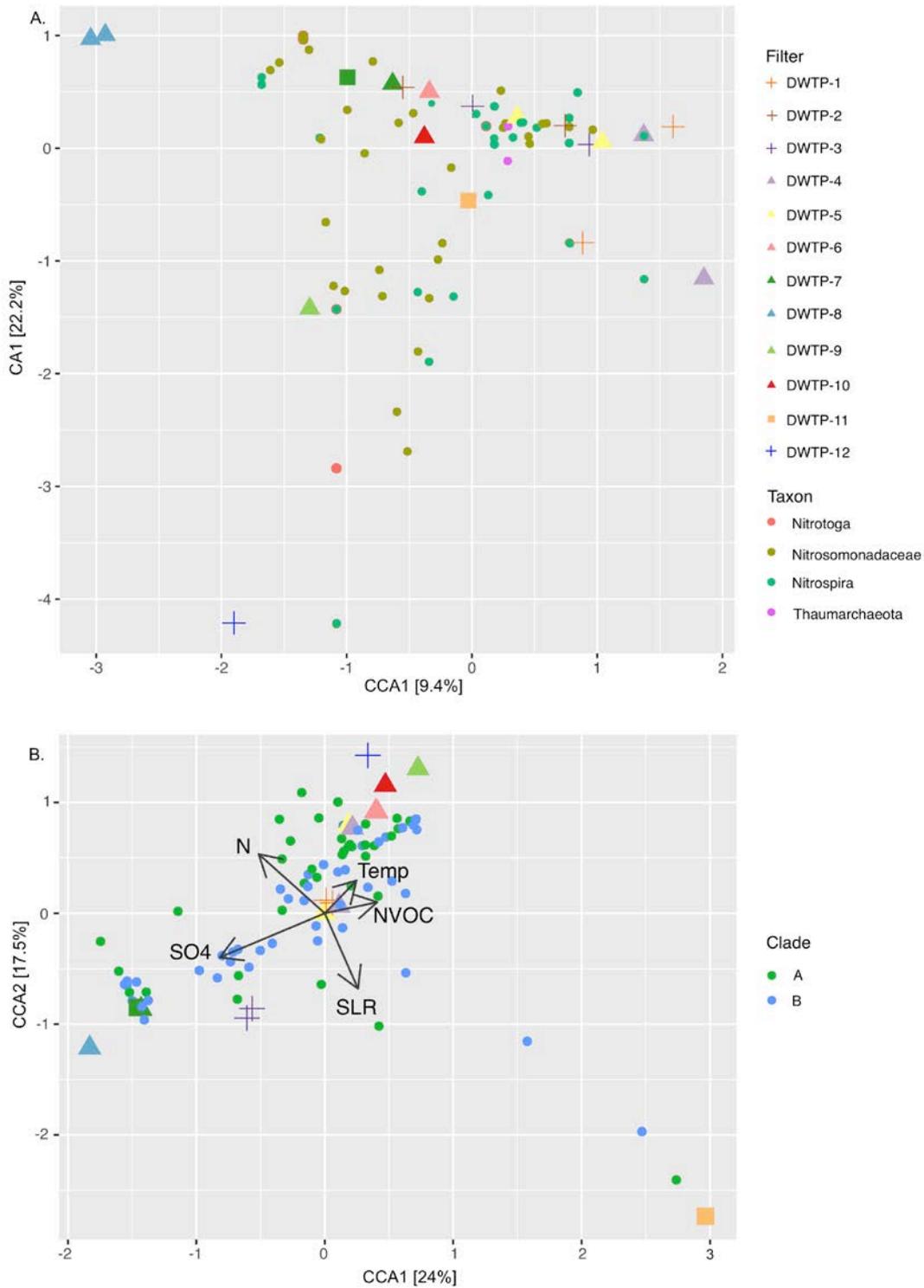


Figure 5. Constrained correspondence analysis of microbial communities based on a) nitrifying communities based on 16S rRNA sequencing (CCA1 and CA1). CCA1 is constrained by temperature, increasing from left to right. b) *Nitrospira amoA* sequencing (CCA1 and CCA2)

Table 1. Primers and thermocycling protocols used in this study

Gene target	Primer name	Sequence	Ta (°C)	Reference
16S Bacteria¹ (Sequencing)	Bakt 341F Bakt 805R	CCTAYGGGRBGCASCAG GGACTACNNGGGTATCTAAT	56	(Yu <i>et al.</i> , 2005)
16S Bacteria¹ (qPCR)	1055F 1392R	ATGGCTGTCGTCAGCT ACGGGCGGTGTGTAC	55	(Ferris and Muyzer, 1996) (Lane, 1991)
16S AOB¹	CTO189FA/B CTO189FC RT1R	GGAGRAAAGCAGGGGATCG GGAGGAAAGTAGGGGATCG CGTCCTCTCAGACCARCTACTG	60	(Kowalchuk <i>et al.</i> , 1997) (Hermansson and Lindgren, 2001)
16S Nitrospira¹	Nspra675F Nspra746R	GCGGTGAAATGCGTAGAKATCG TCAGCGTCAGRWAYGTTCCAGAG	64	(Graham <i>et al.</i> , 2007)
amoA AOA³	CrenamoA23F CrenamoA616R	ATGGTCTGGCTWAGACG GCCATCCATCTGTATGTCCA	55	(Tourna <i>et al.</i> , 2008)
amoA Nitrospira²	Ntsp-amoA 162F Ntsp-amoA 359R	GGATTTCTGGNTSGATTGGA WAGTTNGACCACCASTACCA	48	This study
nxB Nitrospira⁴	nxB169F nxB638R	TACATGTGGTGGAAACA CGGTTCTGGTCRATCA	56.2	(Pester <i>et al.</i> , 2014)
nxB Nitrobacter⁵	NxB-1R NxB-1F	CCGTGCTGTTGAYCTCGTTGA ACGTGGAGACCAAGCCGGG	55	(Vanparys <i>et al.</i> , 2007)

¹ For thermocycling protocol see ¹Gulay *et al.*, 2016; ⁴Pester *et al.*, 2014, ³Tourna *et al.*, 2008

705 ²Thermocycling protocol: 94°C 5:00, 40 cycles of 94°C 30s, 48°C 30s, 72°C 1:00, melt curve (70°C -95°C) 0.2°C/s gradient

⁵Thermocycling protocol: 95°C 10:00, 40 cycles of 95°C 1:00, Ta 1:00, 72°C 2:00, melt curve (70°C -95°C) 0.2°C/s gradient

710

Table 2. Variability explained (fractional) and significance of filter operational parameters and physicochemical features of water and filter material on microbial community structure

Variable	16S rRNA Nitrifiers		<i>amoA Nitrospira</i>	
	Variation explained	<i>p</i>	Variation explained	<i>p</i>

Sulfate (water)	0.060	0.284	0.205	0.001*
Ca (filter material)	0.035	0.730	0.14	0.008
Surface ammonium loading rate	0.024	0.926	0.137	0.016*
N (filter material)	0.093	0.053	0.134	0.017*
Ammonium (water)	0.033	0.833	0.121	0.039
NVOC	0.027	0.914	0.115	0.042*
Temperature	0.094	0.025*	0.104	0.065*
Ca (water)	0.082	0.093	0.099	0.105
Fe (filter material)	0.036	0.724	0.097	0.117
O (filter material)	0.035	0.794	0.082	0.192
DO	0.037	0.665	0.073	0.265
C (filter material)	0.031	0.847	0.072	0.321
Fe (water)	0.038	0.684	0.067	0.392
Mn (filter material)	0.029	0.857	0.066	0.423
P (filter material)	0.059	0.322	0.059	0.504
Cu (water)	0.027	0.858	0.056	0.522
Mn (water)	0.052	0.389	0.037	0.700
Nitrite (water)	0.068	0.341	0.038	0.718
Si (filter material)	0.026	0.842	0.037	0.734
Phosphorus (water)	0.069	0.206	0.037	0.820
Flow rate	0.040	0.683	0.018	0.988
<hr/>				
Total inertia	2.2548		1.6449	
Inertia explained by all variables	2.2437		1.6434	
Inertia explained by significant variables	0.211		1.2905	
Inertia explained by model	0.211		1.101	

*parameters included in the ordination model

715 Ca (filter material) and NH₄ were not included in the model due to high co-linearity with other variables (Figure S4)

Supplementary Information

Figure S1. Representative standard curve (from 10^1 to 10^8 copies) from *amoA Nitrospira*

720 qPCR

Figure S2. Melting curves of *amoA Nitrospira* qPCR product from DWTP-1 (high abundance), DWTP-8 (medium abundance) and standard (from 10^1 to 10^8 copies)

Figure S3. Representative *amoA Nitrospira* PCR products visualized on a 0.7% agarose gel

Figure S4. Phylogenetic tree of sequences used to assess primer coverage

725 Figure S5. Linear regression of a) AOP and NOB and b) comammox and NOB

Figure S6. Spearman correlation matrix of explanatory variables used in ordination

analysis

Figure S7. Pearson correlation matrix of physicochemical/ operational variables with qPCR

data

730 Figure S8. Sample X-ray photoelectron spectroscopy (XPS) analysis of the filter material from the DWTP-5 indicative of the consistent peak patterns observed. a) Complete scan with the primary elements identified; b) Carbon C1s high-resolution scan; c) Nitrogen N1s high-resolution scan; and d) Oxygen O1s high-resolution scan. Of particular importance are the peaks found at 284.3, 286.5, 288.0, 400.0, 400.8, and 531.6 (labelled within the plots) that are assumed to correspond to the presence of proteins in the filter material.

735

Figure S9. Geographic locations of the investigated waterworks

Figure S10. Constrained correspondence analysis of microbial communities based on a) 16S rRNA Bacteria (CA2 and CA3) and b) *amoA Nitrospira* sequencing, CCA3 and CCA4 and d) CCA5 and CA1

740

Table S1. Similarity of comammox classified *Nitrospira nxrB* gene ASVs to closest reference sequences

Table S2. Classification of 16S *Nitrospira* amplicons to comammox reference sequences

Table S3. Physicochemical and operational parameters of the investigated rapid sand

745 **filters**

Table S4. Linear regression of physicochemical/ operational parameters with qPCR data for significant ($p \leq 0.01$) Pearson correlations

Table S5. Details of amplicon sequencing datasets