Metagenomic profiling of ammonia- and methane-oxidizing microorganisms in two sequential rapid sand filters

Lianna Poghosyan, Hanna Koch, Jeroen Frank, Maartje A.H.J. van Kessel, Geert Cremers, Theo van Alen, Mike S.M. Jetten, Huub J.M. Op den Camp, Sebastian Lücker*

Department of Microbiology, Radboud University, Heyendaalseweg 135, Nijmegen 6525AJ, the Netherlands

ARTICLE INFO

Article history:
Received 20 May 2020
Revised 17 July 2020
Accepted 9 August 2020
Available online 10 August 2020

Keywords:
Sand filtration
Nitrification
Comammox Nitrospira
Methanotrophic bacteria
Metagenomics

ABSTRACT

Elevated concentrations of ammonium and methane in groundwater are often associated with microbiological, chemical and sanitary problems during drinking water production and distribution. To avoid their accumulation, raw water in the Netherlands and many other countries is purified by sand filtration. These drinking water filtration systems select for microbial communities that mediate the biodegradation of organic and inorganic compounds. In this study, the top layers and wall biofilm of a Dutch drinking water treatment plant (DWTP) were sampled from the filtration units of the plant over three years. We used high-throughput sequencing in combination with differential coverage and sequence composition-based binning to recover 56 near-complete metagenome-assembled genomes (MAGs) with an estimated completion of ≥70% and with ≤10% redundancy. These MAGs were used to characterize the microbial communities involved in the conversion of ammonium and methane. The methanotrophic microbial communities colonizing the wall biofilm (WB) and the granular material of the primary rapid sand filter (P-RSF) were dominated by members of the Methyloccoccaceae and Methylophilaceae. The abundance of these bacteria drastically decreased in the secondary rapid sand filter (S-RSF) samples. In all samples, complete ammonia-oxidizing (comammox) Nitrospira were the most abundant nitrifying guild. Clade A comammox Nitrospira dominated the P-RSF, while clade B was most abundant in WB and S-RSF, where ammonium concentrations were much lower. In conclusion, the knowledge obtained in this study contributes to understanding the role of microorganisms in the removal of carbon and nitrogen compounds during drinking water production. We furthermore found that drinking water treatment plants represent valuable model systems to study microbial community function and interaction.

© 2020 The Author(s). Published by Elsevier Ltd.
This is an open access article under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

1. Introduction

About 97% of all available water on earth is saline. The remaining 3% is freshwater, of which more than two-thirds is frozen in ice sheets. Thus, only a small fraction of the global freshwater exists as ground and surface water that is available for drinking water production. According to the European Commission (EC, 2016), about 50% of drinking water in Europe is produced from groundwater and 37% from surface water.

Groundwater has a relatively constant composition and may contain high concentrations of iron (Fe²⁺: ≤7.8 mg/L), manganese (Mn²⁺: ≤0.56 mg/L), ammonium (NH₄⁺: ≤0.5 mg/L), and some organic compounds such as methane (CH₄: ≤37 mg/L) (Albers et al., 2015; Li and Carlson, 2014; Osborn et al., 2011). Elevated concentrations of these compounds in groundwater can cause severe problems during drinking water production and distribution, such as excessive biofilm growth, clogging of the biofilter material, pH decrease, pipe corrosion, and taste and odor issues (Beech and Sunner, 2004; Okoniewska et al., 2017; Otten et al., 2016; Sharma et al., 2005; Streese and Stegmann, 2003). To prevent microbial growth in the distribution systems, water treatment plants often use chemical disinfectants such as chlorine, ozone, chlorine dioxide, and chloramines. The use of disinfectants, however, leads to the formation of harmful disinfectant byproducts (Li and Mitch, 2018; Sedlak and von Gunten, 2011) and selection for antibiotic resistant bacteria (Shi et al., 2013; Zhang et al., 2019). In some European countries (e.g., the Netherlands, Denmark, and Switzerland), biofiltration (e.g., rapid (RF) or slow (SF) sand filtration, granular activated carbon filters) are widely applied methods for the removal of the contaminants without the use of disinfectants (Bertelli et al., 2018). Biofilters harbor complex microbial
communities that are introduced via the source water (Yang et al., 2016) and are shaped by the configuration of the treatment process (Li et al., 2017; Pinto et al., 2012). In the filtration, units microbial growth is stimulated on filter material, mediating the biodegradation of organic and inorganic compounds (Proctor and Hammes, 2015). Gases such as methane, hydrogen sulfide, carbon dioxide, and other volatile compounds are removed from the groundwater through gas exchange systems (Trussell et al., 2012). The increased dissolved oxygen in the water caused by this mechanical aeration step serves as an electron acceptor in microbially mediated oxidative reactions, which may ensure the near-complete nutrient removal in the biologically active sand filters.

Naturally occurring ammonium concentrations in deep groundwater are below 0.2 mg/L (WHO, 2003). However, excessive fertilization can lead to increased ammonium concentrations and contamination of groundwater (Kabala et al., 2017). In engineered systems, such as drinking water treatment plants (DWTP), ammonium removal is achieved by the activity of nitrifying microorganisms that oxidize ammonia to nitrate via a series of intermediates. While canonical nitrifying guilds perform ammonia- and nitrite-oxidation in a tight interplay, complete ammonia-oxidizing (co-mammox) bacteria of the genus *Nitrosira* possess all proteins necessary to perform nitrification on their own (Daims et al., 2015; van Kessel et al., 2015). Nitrifying microbial communities of rapid sand filters have been studied before and seem to be represented by different groups of nitrifiers (Albers et al., 2015; Fowler et al., 2018; Gúlay et al., 2016; Oh et al., 2018; Palomo et al., 2016; Pinto et al., 2015; van der Wiel et al., 2009). In these systems, nitrification can be limited by the availability of vital nutrients for this process, such as phosphate and copper. This can cause incomplete nitrification (de Vet et al., 2012; Wagner et al., 2016), leading to incomplete ammonium removal and/or nitrite accumulation (Wilczak et al., 1996), and microbial after-growth in the distribution network (Rittmann et al., 2012).

Methane is a colorless and odorless gas that usually does not present a health risk in drinking water. However, methane gas is highly flammable and can be explosive at elevated concentrations, and it also can serve as substrate for growth of microorganisms in distribution systems. Most of the methane is removed by the mechanical aeration step and remaining amounts are oxidized by aerobic methanotrophs within the sand beds of the filter units. Correspondingly, a number of studies reported methane-oxidizing bacteria colonizing the granular material of the sand filters (Albers et al., 2015; Gúlay et al., 2016; Palomo et al., 2016). However, in contrast to nitrifying microbial communities, methanotrophs in these engineered systems are traditionally less well studied. Methane is a potent greenhouse gas, with a global warming potential of 34 CO₂ equivalents over 100 years (Stocker et al., 2013) and the removal of most methane from the drinking water via aeration causes methane emissions to the atmosphere, contributing to global warming and climate change (Maksimavicius and Roslev, 2020).

Like in many European countries, groundwater is the primary source (65%) for drinking water production in the Netherlands (Geudens and van Grootveld, 2017). In this study, we used genome-resolved metagenomics and gene-centric approaches to analyze microbial communities in rapid sand filters of a Dutch DWTP, with a special focus on ammonia- and methane-oxidizing microorganisms. The groundwater entering this DWTP contains elevated concentrations of organic (CH₄, 5.2 mg/L) and inorganic (NH₄⁺, 0.64 mg/L; Fe²⁺, 8.5 mg/L; Mn²⁺, 0.16 mg/L) nutrients, which are removed in two sequential rapid sand filters. Thus, this DWTP represents an interesting model system to study microbial communities involved in the conversion and removal of these compounds. Metagenomic analysis of samples taken at different stages within this DWTP revealed the key microorganisms involved in ammonia- and methane-oxidation, including novel methanotrophic bacteria from the *Methylphilaceae* family, which were assumed previously to comprise only methlyotrophic bacteria.

## 2. Materials and methods

### 2.1. Sample collection and trace element analysis

Samples were obtained from the pumping station Breehei, a drinking water treatment facility located in Venray, the Netherlands (51°28′54.6″N; 5°59′10.2″E), operated by NV Waterleiding Maatschappij Limburg, which produces drinking water from groundwater. The Breehei DWTP is situated in a rural area with forest (nature reserve) and agricultural land. The raw water entering the DWTP is circumneutral (pH 6.8–7.3) with a temperature of 10 °C. Groundwater is extracted at a depth of 60 m and sprinkled on the saturated filter with a surface of 12.7 m² and bed thickness of 1 m. The surface load over the filter is 3 m³ (m² h)⁻¹. Each filter consists of sand with a grain size of 1.4 to 2.0 mm. At two time points (June 2016 and September 2018) single samples were collected from the top layers of the primary and secondary rapid sand filter (P-RSF and S-RSF) in 50 ml sterile falcon tubes. One sample each from the biofilm formed on the walls of the primary sand filter (WB) was taken in June 2016, May 2017 and September 2018 (Fig. 1). All samples were transferred to the laboratory within 4 h, and were stored at 4 °C for further analysis. Water quality parameters were determined by Aqualab Zuid (https://www.aqualab.nl) over the period 2000–2016 (Table 1).

### 2.2. Methane uptake

Methane-oxidizing capacity was determined for the P-RSF and WB samples collected in 2016. For the P-RSF sample, 2.5, 5, 10, and 20 g of sand material were mixed with 20 ml top water. After settling of the sand, overlaying water samples (20 ml) were transferred into 120 ml serum bottles with ±14.5 mg/L CH₄ in the headspace and incubated at room temperature in a shaking incubator (200 rpm). For the WB sample, the incubations were conducted in triplicates using 1 and 2.5 g of biomass. CH₄ concentrations were determined by sampling 300 µL headspace, which were injected in triplicates into a HP 5890 gas chromatograph (Hewlett Packard, Palo Alto, CA).

### 2.3. DNA extractions

DNA from samples for Illumina sequencing was extracted using two different methods to obtain differential abundance information. DNA from 2016 samples was extracted with DNeasy Blood & Tissue Kit (Qiagen Ltd., West Sussex, United Kingdom) and PowellSoil DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA). Samples (0.5 g) were mechanically disrupted using a TissueLyser (Qiagen) for 2 × 30 s at 30 Hz, followed by DNA extractions according to the manufacturer’s instructions. For the extraction of DNA from samples collected in 2017 and 2018, the DNeasy Blood & Tissue kit was replaced by ammonium acetate (Kovalchuk et al., 2004) and CTAB (Zhou et al., 1996) based extraction methods, respectively.

For long-read Nanopore sequencing, DNA was extracted from P-RSF samples collected in 2016 and 2018 using the CTAB-based extraction method. To avoid shearing of genomic DNA, all bead-beating and vortexing steps were replaced by carefully inverting the tubes several times, and all the pipetting steps were performed using cut-off pipette tips. Genomic DNA was purified twice by phenol:chloroform:isoamyl alcohol (25:24:1) phase extraction (Zhou et al., 1996). Extracted DNA was resuspended in nuclease-free water and stored at 4 °C.
2.4. Illumina library preparation and sequencing

For Illumina library preparation, the Nextera XT kit (Illumina, San Diego, CA, USA) was used according to the manufacturer’s instructions. Enzymatic tagmentation was performed using 1 ng of DNA per sample, followed by incorporation of the indexed adapters and library amplification. After subsequent purification using AMPure XP beads (Beckman Coulter, Indianapolis, IN, USA), libraries were checked for quality and size distribution using the 2100 Bioanalyzer with the High Sensitivity DNA kit (Agilent, Santa Clara, CA, USA). Library quantitation was performed by Qubit using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). After dilution to 4 nM final concentration, the libraries were pooled, denatured, and sequenced on an Illumina MiSeq. Paired-end sequencing of 2 × 300 base pairs was performed using the Illumina MiSeq Reagent Kit v3 according to the manufacturer’s protocol.

2.5. Nanopore library preparation and sequencing

The DNA Library preparation for Nanopore sequencing was performed using the Ligation Sequencing Kit 1D (SQK-LSK108) in combination with the Native Barcoding Expansion Kit (EXP-NBD103 or EXP-NBD104) according to the manufacturer’s protocol (Oxford Nanopore Technologies, Oxford, UK). For the detailed protocol, see Supplementary Methods. The libraries were loaded on a Flow Cell (R9.41) and run on a MinION device (Oxford Nanopore Technologies, Oxford, UK), according to manufacturer’s instructions. Base calling after sequencing was done using Albacore v2.1.10, (Oxford Nanopore Technologies) for the 2016 and guppy_basecaller in combination with gGuppy (Oxford Nanopore Technologies; Version 2.3.7+e041753) for the 2018 sample.

2.6. Assembly

Raw Illumina sequence data were processed for quality-trimming, adapter removal, and contamination-filtering, using BBduk (BBTOOLS v37.17: http://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/bbduk-guide/). All sequencing reads from the same sampling location were co-assembled using metaSPAdes v3.10.1 (Nurk et al., 2017) with the following parameters: k-mer sizes 21, 33, 55, 77, 99 and 127, minimum contig length 1500 base pairs (bp). Raw Nanopore reads were quality trimmed using Filtlong v0.2.0 (https://github.com/rrwick/Filtlong) with minimum read length 1000 bp and error rate <20%. Porechop v0.2.4 (https://github.com/rrwick/Porechop) was used to remove adapters and split chimeric reads with default settings. Trimmed reads were assembled using Canu v1.8 (Koren et al., 2017) with minimum read length 1000 bp, corrected read error rate 0.105, and genome size 5 Mb. To correct error rates, the trimmed Nanopore reads were mapped to the assembly using Minimap2 v2.16-r922 (Li, 2018), which then was polished with Racon v1.3.1 (Vaser et al., 2017). Subsequently, Racon v1.3.1 was used to further polish the assembly twice with Illumina reads obtained from the 2018 P-RSF sample.

2.7. Metagenome binning

Differential coverage information was determined by separately mapping the sequencing reads from each sample and DNA extraction method against the obtained co-assemblies, using Burrows-Wheeler Aligner v0.7.15 (BWA) (Li and Durbin, 2010) and employing the “mem” algorithm. For the long-read Canu assembly, only the Illumina reads from P-RSF 2018 samples were mapped. The generated sequence alignment map (SAM) files were converted to binary format (BAM) using SAMtools (Li et al., 2009). Metagenome binning was performed using anvio v5.3, followed by manual bin refinement using the anvio interactive interface (Delmont and Eren, 2016; Eren et al., 2015). For details, see Supplementary Methods. Completeness and contamination (referred to as redundancy in this study) of bins was assessed by CheckM v1.0111 (Parks et al., 2015). Based on the suggested standards (Bowers et al., 2017), the bins were defined as high-quality (≥90% completeness and ≤5% redundancy), complete small subunit rRNA operon, ≥18 tRNAs, and medium-quality (≥70% complete and ≤10% redundancy) metagenome-assembled genomes (MAGs). MAGs were dereplicated using dRep v2.2.3 (Olmi et al., 2017) at 99% average nucleotide identity (ANI) used for clustering. Within each cluster, the best MAG was selected based on completeness, redundancy, N50 of contigs, and fragmentation. GTDB-Tk v0.3.2 (Parks et al., 2018) was used for taxonomic assignment of the final MAGs. Phyla are named according to the recently suggested nomenclature (Whitman et al., 2018) using standardized phylum suffix -ota.
2.8. Abundance estimation of MAGs

To calculate the relative abundance of the dereplicated MAGs in each sample, reads from all samples were individually mapped to each co-assembly using BWA v0.7.15 (Li and Durbin, 2010) as described above. The coverage of each MAG was calculated using CheckM v1.01.11 (minimum alignment length 0.95) (Parks et al., 2015) and was normalized by multiplying this coverage with a normalization factor (sequencing depth of the largest sample divided by the sequencing depth of each individual sample). The distribution of MAGs was calculated as percentages by dividing a MAG’s coverage in each sample by the total coverage of the respective MAG in all samples.

2.9. Functional analysis

For the gene-centric approach, co-assembly across all samples was performed using MEGAHIT v1.1.1–2 (Li et al., 2015). Open reading frames (ORFs) in the new co-assembly and the MAGs obtained above were predicted using Prodigal v2.6.3 (Hyatt, 2010), which was set to include partial ORFs. Custom-build hidden Markov models (HMMs) (Eddy, 2011) of specific marker proteins were used (Supplementary Methods) to annotate all ORFs using hmmscan (HMMER v3.1b2; http://hmmer.org). The HMM for RNA polymerase subunit beta (RpoB) was downloaded from FunGene (Fish et al., 2013). Remaining ORFs in the MAGs were annotated using Prokka v1.12-beta (Seemann, 2014). The annotations of all genes discussed in this study were confirmed by BLAST against the TrEMBL, SwissProt and NCBI nr databases. Subcellular localization of the proteins was predicted by SignalP 5.0 (Armenteros et al., 2019) and TMHMM 2.0 (Krogh et al., 2001).

Functional gene-based abundances of ammonia- and methane-oxidizing microorganisms were estimated using competitive metagenomic read recruitment to ensure unique mapping. For this, reads from each metagenomic sample were mapped using bowtie2 v2.3.1 (Langmead and Salzberg, 2012) in ‘-very-sensitive’ mode against extracted partial and complete sequences of rpoB and the ammonia as well as the particulate and soluble methane monoxygenase subunit A genes (amoA, pmoA and mmoX, respectively). SAMtools v1.6 (Li et al., 2009) was used to obtain the number of mapped reads. Reads per kilo base per million mapped reads (RPKM)-values were used to correct for differences in sequencing depth and gene length. To estimate the relative abundance of microorganisms encoding ammonia and methane monoxygenases in each sample, the normalized read counts were calculated as fraction of the normalized read counts of the identified rpoB genes.

2.10. Phylogenomic and phylogenetic analyses

The up-to-date bacterial core gene (UBCG) pipeline (Na et al., 2018) with default parameters was used to extract and concatenate bacterial core gene sets. To infer the phylogeny of the archaean MAGs, the anv’i’o phylogenomic workflow (http://merenlab.org/2017/06/07/phylogenomics) was used to individually align and concatenate 162 archaeal single-copy genes. Maximum-likelihood trees were calculated using RAxML version 8.2.10 (Stamatakis, 2014) on the CIPRES science gateway (Miller et al., 2010). For details, see Supplementary Methods. Amino acid sequences of type II DMSO reductase-family enzymes, and of ammonia and methane monoxygenases were aligned using ARB v5.5 (Ludwig et al., 2004). The maximum-likelihood trees were calculated using RAxML HPC-HYBRID v8.2.12 on the CIPRES and IQ-tree webserver (Trifinopoulos et al., 2016) as described in Supplementary Methods. All phylogenetic trees were visualized in iTOL (Letunic and Bork, 2016).

### Table 1

| Water quality parameter | Groundwater | Draining water | Unit (mg L⁻¹) |
|-------------------------|-------------|---------------|---------------|
| **Average**             | 7.00        | 7.00          | pH            |
| **Range**               | 6.8–7.3     | 7.00–7.7      | **Average**    |
| **Groundwater**         | 7.00 (n = 46) | 7.00 (n = 3)  | **Draining water** |
| **Draining water**      | 7.00 (n = 46) | 7.00 (n = 3)  | **Unit (mg L⁻¹)** |

**Not applicable or were not analysed. No standards.**

| **Water quality parameter** | **Groundwater** | **Draining water** | **Range** | **Average** |
|-----------------------------|-----------------|-------------------|-----------|-------------|
| **Average**                 | 7.00 (n = 46)   | 7.00 (n = 3)      | **Draining water** |
| **Draining water**          | 7.00 (n = 46)   | 7.00 (n = 3)      | **Range**   |
| **Unit (mg L⁻¹)**           | **Average**     | **Draining water** |

**Not applicable or were not analysed. No standards.**
2.11. Data visualization

Manuscript figures were generated using ggplot2 (Wickham, 2016), Rstudio (Racine, 2012) and the anv/ö interactive interface (http://merenlab.org/2016/02/27/the-anvio-interactive-interface).

3. Results

3.1. Recovery of metagenome-assembled genomes

Over the period 2016 to 2018 a total of 7 samples from DWTP Breehei were collected (Fig. 1) and sequenced. This resulted in a total of 125 million paired-end Illumina sequencing reads. For each sample location more than 70% of the respective reads could be co-assembled into ~413, 184, and 249 Mb sequencing data for WB, P-RSF, and S-RSF, respectively (Table S1). In addition, P-RSF samples were also sequenced using the Oxford Nanopore long-read platform to improve the assembly of the most abundant microorganisms. Overall, binning of 4 individual metagenome assemblies based on sequence composition and differential coverage patterns resulted in 78 near-complete Illumina and 7 Nanopore MAGs (Table S2). All MAGs obtained from Illumina co-assemblies as well as the Nanopore assembly were dereplicated at strain level (99% ANI), which yielded 50 medium and 6 high-quality (MiMAG standards; Bowers et al., 2017) non-redundant MAGs (Table 2, Table S2) that were used for downstream analyses. Given the high number of single nucleotide variants observed during bin refinement using the anvö interactive interface, 15 MAGs were categorized as population-level genomes (Table 2). All MAGs were classified at the lowest possible taxonomic level using GTDB-tk (Parks et al., 2018), indicating affiliation with 1 archaeal and 12 different bacterial phyla (Fig. 2, Table 2).

3.2. Distribution and taxonomic composition of the DWTP microbiome

The influence of sampling location, sampling time, and DNA extraction method on the distribution of recovered MAGs was analyzed by hierarchical clustering using Euclidean distance metrics with a Ward linkage algorithm, allowing grouping of the MAGs based on their occurrence patterns in the different samples (Fig. 2).

Overall, the choice of a DNA extraction method had no pronounced influence on the distribution of MAGs across samples, except for the sample from P-RSF in 2016. In this specific sample, the normalized coverage values indicated a strong extraction bias for DNA extracted using the Power soil (P-RSF16_PS) compared to the Blood and Tissue kit (P-RSF16_BT; Table S2). Notably, this bias was not observed for the 2016 S-RSF samples extracted with the same kits, or for any other DNA extraction performed using the Power soil kit. It is thus difficult to conceive that the extraction method was affecting specific members only in the 2016 P-RSF microbial community, and the reason for the observed bias remains unclear. Overall, the sampling location had the most substantial effect on microbial diversity and abundance recovered in our study. While some MAGs obtained from WB and P-RSF were also present in S-RSF, the overall community of the S-RSF community clearly differed from the other sampling locations (Fig. 2). The microbial communities of WB and P-RSF were generally more similar, but sampling location and time influenced the relative abundance of MAGs across all samples.

To gain insights into the overall microbial community structure and diversity of the DWTP Breehei, 16S rRNA gene sequences were retrieved directly from the metagenomic assemblies. Both full-length and partial 16S rRNA gene sequences were used for further analyses, since only 21 to 32% of 16S rRNA reads were fully assembled (Figure S2). Subsequently, microbial community composition was analyzed at the phylum and family levels (Figure S3). Although changes in abundance were observed, phylum level classification revealed no differences in microbial community composition between the sampling locations (Figure S3A). At the family level, the S-RSF microbial community was clearly distinct from the WB and P-RSF community (Figure S3B), corroborating the MAG occurrence-pattern-based observations when samples were organized using Euclidian distance and Ward ordination (Fig. 2).

16S rRNA gene-based taxonomic profiling identified 14 phyla with relative abundances ≥1% (Figure S3A). Gammaproteobacteria (17–67%), Bacteroidota (5–17%), Acidobacteria (2–15%), Alphaproteobacteria (4–12%), Planctomycetota (2–10%) “Ca. Patescibacteria” (CPR; 1–7%), and Nitrospirota (2–6%) were the most dominant phyla identified in all samples (Figure S3A). In most cases, this 16S rRNA gene-based analysis corroborated the community structure found based on the taxonomic affiliation of the recovered MAGs. Additionally, one MAG belonged to the candidatus phylum “Ca. Eisenbacteria” and one Crenarchaeota (Thaumarchaeota) MAG were obtained from WB and S-RSF (Fig. 1), but were of low abundance in the 16S rRNA datasets.

3.3. Genome functional profiling

During the drinking water treatment process, contaminants like iron, manganese, reduced sulfur species, ammonium and methane are removed. However, our understanding of the microbial and geochemical processes contributing to the removal of these compounds during sand filtration is still limited. To learn more about the microorganisms driving the removal of methane and ammonium in this rapid sand filtration system, all recovered MAGs were screened for functional marker proteins involved in nitrification, and methane and other one-carbon (C1) compound oxidation (Fig. 3, Table S3) using custom-build HMM models. Details on microorganisms involved in iron and sulfur cycling are given in the Supplementary Material.

3.3.1. Ammonia and nitrite oxidation

16S rRNA gene sequence analysis revealed the presence of nitifying bacteria affiliated with the families Nitrospiraceae and Nitrosomonadaceae in the DWTP Breehei. Nitrospiraceae dominated the nitifying microbial community in all samples, but their abundance patterns differed along the different sampling locations, with the lowest abundance in P-RSF. In contrast, canonical ammonia-oxidizing bacteria (AOB) affiliated with the Nitrosomonadaceae showed very low abundance in all samples (Figure S3). Metagenomic consensus binning allowed the recovery of five Nitrospira MAGs from WB and P-RSF and one MAG of an ammonia-oxidizing archaeon (AOA) affiliated with the genus Nitrosoarchaeum. Despite the detection of Nitrosomonadaceae in WB and S-RSF samples in 16S rRNA-based analyses, no metagenomic bin of this taxonomic group was recovered.

All MAGs were screened for key genes of autotrophic nitrification, including the gene sets encoding ammonia monooxygenase (AMO) and nitrite oxidoreductase (NXR). AmoA is often used as a functional and phylogenetic marker for ammonia-oxidizing microorganisms (Junier et al., 2008; Pester et al., 2012; Pjevac et al., 2017) and thus was used to examine the full ammonia oxidation potential in the DWTP Breehei. Phylogenetic analysis placed our metagenome-derived amoA sequences into five divergent groups (Fig. 4A). The observed variations in the nitifying microbial community structure and abundance between the different samples might be influenced by seasonality, with the largest relative abundance in spring 2016 and the lowest during fall 2018 (Fig. 4B). However, to confirm this trend, more frequent samples across a
longer time span would be required, also considering variations in physicochemical parameters.

Of the five Nitrosira MAGs recovered in this study, one (P-RSF-IL-20) contained amoA genes affiliated with clade A, and two (WB-IL-04, WB-IL-16) with clade B (Fig. 3A, Table S3). In protein-based phylogenetic analyses, most of the clade A and B community amoA sequences most closely clustered with sequences derived from DWTP metagenomes (Figure S4). Clade A community amoA genes were detected in all samples from the DWTP Breehei (0.1–3.6% of the total community) and were the most abundant amoA type in P-RSF. The clade A comammox Nitrosira MAG (P-RSF-IL-20) had the highest coverage in the 2016 P-RSF sample extracted with the Blood & Tissue Kit, while this coverage decreased drastically with another DNA extraction method, indicating a strong extraction bias as discussed above (Table S2). Consequently, the average coverage of this MAG was higher in S-RSF than in P-RSF. In contrast, clade B community amoA recruited the largest number of all assembled reads in WB (0.8–5.9%) and S-RSF samples (1.7–4.1%; Fig. 6A), but none of the two clade B-affiliated MAGs was detected in S-RSF, suggesting that not all clade B co-
mammox Nitrosospira genomes were recovered. The low abundance of clade B combammmox amoA in P-RSF (Fig. 4B) indicates an adaptation of clade B combammmox Nitrosospira to specific niches. Except for one MAG (P-RSF-IL-22), all Nitrospira MAGs contained a nxrAB gene cluster encoding for the alpha and beta subunit of the NXR complex (Fig. 3A, Table S3). MAG P-RSF-IL-22 was of medium quality (74.4% estimated completeness; Table 2) and did not contain any of the genes required for nitrification (Table S3). Consistent with previous studies (Palomo et al., 2019; Poghosyan et al., 2019), phylogenomic analysis using a concatenated alignment of 91 single-copy core genes showed that clade A and B combammmox Nitrosospira formed monophyletic clades within Nitrospira lineage II, and all amoA-containing MAGs were correctly affiliated with their respective combammmox clade (Fig. 5A).

One MAG (WB-IL-22) affiliated with the genus Nitrosoarchaeum (AOA; Fig. 5B) was retrieved from the WB metagenome; however, it lacked an amoA sequence. Since a contig (<1200 bp) containing a Nitrosoarchaeum-like amoA was identified in the metagenome, the gene most likely is missing from this MAG due to the size cutoff used during binning (1500 bp). Consistent with gene-based analyses (0.1–1.7%; Fig. 4B) the Nitrosoarchaeum MAG was found solely in WB, where it accounted for merely ~0.2% of all the assembled reads in the WB datasets (Fig. 2, Table S2). The recovered archaeal amoA sequence had high similarities to Nitrosoarchaeum koreense and sequences derived from metagenomic analyses of diverse habitats (Figure S4). Betaproteobacterial amoA sequences of members of the genera Nitrosomonas and Nitrosospira were identified in all samples, but were of low abundance. In addition to the characterized ammonia-oxidizing clades, recently the unclassified Gammaproteobacteria “MBAE14” genome was found to contain putative AMO genes (Mori et al., 2019). From the DWTP Breehei metagenomes, three putative AmoA sequences clustered in this novel sequence group, albeit with low similarities to the putative enzyme of MBE14 (Fig. 4A, S4).

The NXR belongs to the type II dimethyl sulfoxide (DMSO) reductase enzyme family (Lückner et al., 2010; Meincke et al., 1992), which also contains respiratory nitrate reductases (NARs) (Simon and Klotz, 2013). Consequently, many HMM profiling approaches fail to differentiate the two homologous groups. Therefore, a phylogenetic tree was constructed to classify the catalytic alpha subunits (NxrA/NarG) identified in the DWTP metagenomes (Fig. 6).

Besides the Nitrosira genomes, six additional MAGs contained NXR/NAR gene clusters (Fig. 3A). Phylogenetic analysis revealed
that the sequences derived from alphaproteobacterial MAGs (P-RSF-IL-15, P-RSF-IL-21, S-RSF-IL-04) were closely affiliated with NarG sequences of known nitrate reducers (Figure S5). The sequences from a MAC affiliated with the Pedosphaeraceae (WB-IL-17) and two Planctomycetota MAGs (WB-IL-13, S-RSF-IL-06) shared low similarities with known NXR s, but clustered with Nitrospira, Nitrospina and anammox bacteria (Figure S5). While the catalytic NxrA subunit of WB-IL-13 contains one Fe-S domain typical for DMSO reductase family molybdoproteins (Lücker et al., 2010), this binding motif for Fe-S clusters is absent from the WB-IL-17 and S-RSF-IL-04 NXR-like proteins. Similar to other periplasmic NXRs (Lücker et al., 2013; Lücker et al., 2010; Spieck et al., 1998), these putative NXRs contain N-terminal twin-arginine motifs for protein translocation into the periplasmic space. Putative genes encoding the NXR gamma subunit (nxrC) were found to be co-localized with nxrAB in the WB-IL-13 and WB-IL-17 MAGs. Similar to Nitrospina gracilis (Lücker et al., 2013) and “Ca. Nitrotoga fabula” (Kitzinger et al., 2018), these putative NXRs contain N-terminal signal peptides necessary for translocation via the Sec pathway. The Pedosphaeraceae (WB-IL-17) and one of the Planctomycetota MAGs that affiliated with the Phycisphaerales (WB-IL-13) were present in WB samples, but of low abundance (Fig. 2, Table S1). The second, unclassified Planctomycetota MAG (S-RSF-IL-06) was detected only in S-RSF and accounted for 0.5–2% of all assembled reads in the S-RSF datasets (Table S1). Additionally, some of the metagenome-derived NXR/NAR sequences were placed

Fig. 3. Metabolic potential of the 56 MAGs recovered from DWTP Breehei A. Phylogenetic tree based on the concatenated alignment of 49 ribosomal proteins (Table S3) using the anvio phylogenomics workflow. Presence or absence of genes for ammonia, nitrite, methane, and C1 utilization are indicated by filled or shaded colored boxes, respectively. Gray bars represent estimated genome completeness. The figure was generated using the anvio interactive interface. B. Schematic pathway models for complete nitrification, and methane and C1 oxidation, based on enzyme complexes identified in the metagenome. ALDH, aldehyde dehydrogenase; AMO, ammonia monooxygenase; COX, cytochrome c oxidase; FDH, formate dehydrogenase; FoD, methylenetetrahydrolfolate dehydrogenase/cyclohydrolase; GSH, glutathione-linked formaldehyde oxidation; HAO, hydroxylamine dehydrogenase; H4MPT, tetrahydromethanopterin; HURM, hydroxylamine-ubiquinone redox module; pMMO, particulate methane monooxygenase; sMMO, soluble methane monooxygenase; MDH, lanthanide and calcium-dependent methanol dehydrogenases; NAR, nitrate reductase; NXR, nitrite oxidoreductase.
Fig. 4. Diversity and abundance of ammonia and methane-oxidizing microorganisms in DWTP Breehei. (A) Phylogenetic tree of CuMMO subunit A proteins recovered from the DWTP metagenomes by gene-centric and genome-resolved approaches. Numbers in brackets indicate sequences per metagenome, number of genes in recovered MAGs, and reference datasets for each clade, respectively. Amo, ammonia monooxygenase; Emo, ethane monooxygenase; Pmo/Pxm, particulate methane monooxygenase. Bootstrap support values ≥90% are indicated by black dots; the scale bar indicates estimated amino acid substitutions. Normalized abundances of (B) ammonia and (C) methane oxidizers are shown as proportion of the recovered ammonia and methane monooxygenases, respectively, to the normalized rpoB abundance. The mmoX gene was used to calculate the abundance of sMMO-containing methanotrophs; most mmoX reads were recruited by MAG P-RSF-NP-02 (70–100%). As most available high-quality betaproteobacterial AOB genomes (n = 17, comp. ≥90%, number of contigs <5) contain between 2 and 3 amoA copies, and MAG P-RSF-NP-02 contained 4 mmoX copies, abundances of both genes were normalized for gene copy numbers. Sample labels correspond to Fig. 2.

Fig. 5. Maximum-likelihood phylogenomic analysis of (A) Nitrospiraceae based on 91 bacterial single-copy core genes and (B) Nitrosopumilaceae based on 162 archaeal proteins. DWTP genomes obtained in this study are shown in red. Bootstrap support values ≥95% are indicated by black dots; the scale bars indicate estimated nucleotide (A) and amino acid (B) substitutions. The positions of the outgroups are indicated by arrows. Environmental origins of the respective genomes are represented by shaded circles. Genome names correspond to NCBI RefSeq entries, while in (B) classifications at higher taxonomic levels (gray shadings) are according to the GTDB database (Table S4). In (A) Nitrospira sublineages are delineated by black bars, the comammox clades are highlighted by turquoise shadings.
into the clade containing the novel NXR type of “Ca. Nitrotoga fabula” (Figure 5B) (Kitzinger et al., 2018).

3.3.2. Methane and one-carbon metabolism

Water analyses indicated that most of the methane present in the incoming groundwater was removed during the aeration step, and the remainder was oxidized in the P-RSF (Table 1). Consequently, we determined microbial methane uptake rates in the P-RSF and WB samples. Complete methane oxidation was achieved in all samples within 5 to 8 days of incubation, except for the control containing only raw water (Figure S1). As expected, methane consumption in the P-RSF sample increased with an increasing amount of sand material, whereas no significant difference in methane uptake was detected between the incubations with 1 and 2.5 g of WB biomass (Figure S1). These identical oxidation rates in the WB incubations seem counterintuitive, but may be caused by sample inhomogeneity and uneven distribution of methanotrophic bacteria within the biofilm.

Taxonomic profiling of the extracted 16S rRNA genes (Figure S3) identified Gammmaproteobacteria as the most dominant taxa in WB (40–55%) and P-RSF (61–67%) samples. On the family level, Methylophilaceae and Methylibacteriaceae formed the most abundant groups in WB and P-RSF (Figure S3B). In total, 14 MAGs belonging to these two phylogenetic groups were recovered, which were mainly present in WB and P-RSF samples (Figure 2).

Phylogenomic analysis of the 14 Methylophilaceae and Methylibacteriaceae-affiliated MAGs using a concatenated alignment of 92 single-copy core genes showed that three of the seven Methylophilaceae MAGs were affiliated with the genus Methylobacterium (Figure 7A). Of the remaining MAGs, two (P-RSF-IL-13 and WB-IL-06) were distantly related to Methylobacterium and one (WB-IL-19) to Crenothrix, while P-RSF-IL-24 clustered separately from the known genera within this family (Figure 7A). Six out of seven Methylibacteriaceae MAGs were affiliated with the genus Methylophilus, whereas P-RSF-NP-02 formed a separate branch within the Methylibacteriaceae family (Figure 7B).

The relative coverage of the Methylibacteriaceae MAGs was much higher than for the Methylophilaceae (Figure 2, Table S1). These results are consistent with 16S rRNA gene profiling, where Methylophilaceae constituted the most abundant family in WB (6–27%) and P-RSF (28–47%) samples and disappeared in S-RSF samples (Figure S3). Particularly, the MAG P-RSF-NP-02 had extremely high coverage and relative abundance in the WB (1.6–179 ×; 0.1–9.6%) and P-RSF (27–744 ×; 1.9–35.8%) metagenomes (Table S1). Based on 16S rRNA gene similarity, “Ca. Methylosepiumpurum turicensis” MM5–10A–171 (Salcher et al., 2019) is the closest described relative of P-RSF-NP-02 (95.7% sequence identity). This value is below the similarity cutoff for species delineation (Stackebrandt and Ebers, 2006), which, in combination with the distinct clustering in the phylogenomic analysis (Figure 7B), indicates that P-RSF-NP-02 probably belongs to a novel genus within the Methylibacteriaceae. The estimated genomic average nucleotide identity value (ANI; 77.26%) also indicated a novel species distinct from “Ca. M. turicensis”. The Breehei DWTP metagenomes and the recovered MAGs furthermore were screened for genes encoding the soluble (sMMO) or particulate (pMMO) methane monooxygenases (Table S3). Notably, a complete operon encoding sMMO (mmoX1X2Y2Z2DC) was identified in the highly-covered Methylibacteriaceae MAG P-RSF-NP-02 (Figure 3, Table S3). This was surprising as no previously described member of the family Methylibacteriaceae contains any methane monooxygenase (Salcher et al., 2019). P-RSF-NP-02 harbors two additional copies of the mmoX gene encoding the alpha subunit of the sMMO (Table S3). The two orphan mmoX copies share high sequence identities on amino acid level with the operon mmoX1 (99.8–100%), while mmoX2 has ~97% identity to mmoX1 and the two orphan mmoX. Within the NCBI RefSeq database, the most similar proteins to the MmoX copies of P-RSF-NP-02 belong to Methylibaculum buryatense with similarities of 83–84% (Kaluzhnaya et al., 2001). The abundance of methanotrophs in the total community was estimated based on pmoA, pxaA and mmoX gene coverages (Figure 4C). The sMMO-containing methanotrophs were present at high abundances in the 2018 WB (6–11%) and P-RSF (2.5–56%) samples (Figure 4C), where MAG P-RSF-NP-02 also dominated the total microbial communities (Figure 2). The remaining samples were dominated by pMMO-containing methanotrophs (Figure 4C). The gene cluster encoding the structural pMMO subunits (pmoABC) was identified in all Methylibacteriaceae genomes. Furthermore, five MAGs also harbored the highly divergent pxaABC gene cluster encoding for pXMO (Figure 3) (Tavormina et al., 2011). Phylogenetic analysis of the pMMO alpha subunits (PmoA/PxmA) revealed that all sequences recovered in this study, including the metagenome-derived ones, were affiliated with gammadimotrichial methanotrophs (Figure S4). Besides pMMO, the three Methylophilaceae MAGs (WB-IL-19, P-RSF-IL-13, P-RSF-IL-24) also possessed sMMO (Figure 3, Table S3).

In both methanotrophs and non-methanotrophic methylotrophs, methanol oxidation to formaldehyde is catalyzed by pyrroloquinoline quinone-dependent (PQQ) methanol dehydroge-
nases (MDH) (Keltjens et al., 2014). In this study, the rare earth element-dependent MDH (Ln-MDH) was detected in all Methylomonadaceae and Methylphilaceae, in one Deltaproteobacteria (S-RSF-IL-11), and one Hypyhlocyrum (P-RSF-OP-07) MAG. In contrast, the calcium-dependent MDH (Ca-MDH) was found only in few genomes (Fig. 3A, Table S2). Thus, all analyzed methanotrophs and methyloths can oxidize methanol either to formaldehyde by Ca-MDH, or possibly directly to formate by Ln-MDH (Fig. 3B) (Pol et al., 2014). Four formaldehyde-oxidizing pathways were identified, including oxidation by single aldehyde dehydrogenases, as well as via tetrahydromethylredoxin (H4MPT), tetrahydrofumurate (H4F), or glutathione (GSH)-linked pathways (Fig. 3). All the MAGs contained formate dehydrogenases (FDHs), necessary for the oxidation of formate to carbon dioxide.

4. Discussion

4.1. DWTP microbiome

The DWTP Breehei microbiome was evaluated over a time period of three years and characterized using genome-resolved and gene-centric metagenomic approaches. The microbial community structure was studied for the filter material of two sequential rapid sand filters and the biofilm formed on the walls of the primary sand filter (Fig. 1). Similar to some Danish DWTPs (Gülay et al., 2016), the microbial community composition of the two sequential sand filters differed substantially. In total, 56 dereplicated near-complete MAGs were recovered, comprising 23, 64, and 14% of the total assembled reads for WB, P-RSF, and S-RSF, respectively (Table S1). These MAGs expand our knowledge on the genomic inventory of the main microorganisms involved in contaminant removal from groundwater to produce drinking water. However, assembly statistics indicated that the obtained metagenomic information especially for WB and S-RSF only covers a part of the diversity. In general, the genome-centric approach is a powerful tool to analyze the functional potential of an environmental sample based on recovered MAGs. However, due to low abundance or strain diversity, it is often difficult to obtain good-quality genomes for many microorganisms (Syrzyba et al., 2017). Thus, to examine the full ammonia and methane oxidation potential, we also used a gene-centric approach (see below). More information on the microorganisms involved in the removal of other contaminants is given in the Supplementary Results and Discussion.

4.2. Ammonia and nitrite metabolism

The removal of ammonium and nitrite is a vital step in drinking water treatment. Although the drinking water produced in DWTP Breehei is of high quality and free of any nitrogen compounds, the ranges of nitrate concentrations in the effluents of the P-RSF and S-RSF indicate intermediate nitrate accumulation, which can be caused by incomplete nitrification (de Vet et al., 2012; Wagner et al., 2016). Thus, profound insights into the microbial key players and their metabolic capabilities and limitations are crucial for optimizing and stabilizing N removal in these systems. The genomic potential for nitrification was mainly identified in MAGs af-
filiated with the genus *Nitrospira*, which however did not display high abundances as observed in other DWTP systems (Albers et al., 2015; Gülay et al., 2016; Palomo et al., 2016). Comammox *Nitrospira* were the most abundant nitrifying guild in all DWTP Breehei samples (Fig. 4B), corroborating previous findings of *Nitrospira* dominating the nitrifying microbial community in DWTPs (Albers et al., 2015; Gülay et al., 2016). Until recently, the dominance of this group was puzzling since *Nitrospira* were always regarded as strictly nitrite-oxidizing bacteria. By now, this imbalance in abundance of *Nitrospira* and canonical AOB can be explained by the presence of complete nitrifying *Nitrospira* in many sand filtration systems (Palomo et al., 2018; Pinto et al., 2015; Wang et al., 2017). Notably, complete nitrifiers are the most abundant nitrifying group in several Danish RSFs (Fowler et al., 2018) and were identified as key drivers of ammonia and nitrite oxidation in these systems (Gülay et al., 2019). Consistent with these previous results, comammox *Nitrospira* also apparently outcompeted canonical AOB and AOA in DWTP Breehei. Comammox clade A was found to dominate in P-RSF samples and clade B in WB and S-RSF (Fig. 4B). Thus, sampling location substantially influenced the abundance of the two comammox clades, indicating potential niche partitioning between them. The habitat preferences of different nitrifiers might be explained by their ammonia oxidation kinetics and substrate affinities (Kits et al., 2017; Martens-Habbena et al., 2009; Prosser and Nicol, 2012). According to the kinetic theory of optimal pathway length (Costa et al., 2006) and physiological analyses of *H. inopinata* (Kits et al., 2017), comammox *Nitrospira* are K-strategists that have a competitive advantage in environments with very low ammonium fluxes. However, since no comammox clade B enrichment culture is currently available, we can only speculate about the niche-defining metabolic capabilities of this group. It has been shown that in Danish groundwater and Chinese surface water-treating sand filters with influent ammonium concentrations ranging from 0.01–0.53 mg-N/L and ~0.01–0.035 mg-N/L, respectively, clade B comammox *Nitrospira* dominated the nitrifying microbial populations (Fowler et al., 2018; Hu et al., 2020). A recent study also reported that nitrification activity in forest and paddy soils when subjected to ammonium limitation is associated with clade B rather than clade A comammox *Nitrospira* (Wang et al., 2019). Comammox clade B appeared to dominate also in forest soil under increasing nitrogen load and decreasing pH (Shi et al., 2018). Under acidic conditions, ammonia (NH₃) is increasingly protonated to ammonium (NH₄⁺), resulting in extremely low concentrations of bioavailable ammonia, the substrate of AMO. The higher abundance of clade B comammox *Nitrospira* in S-RSF compared to P-RSF observed here also suggests their adaptation to ammonium-depleted environments. These results indicate that clade B comammox *Nitrospira* may exhibit an even lower half-saturation constant (Kₘ) and higher substrate affinities than clade A species and thrive at extremely low ammonium concentrations. However, it will be required to obtain clade B comammox *Nitrospira* in culture to ascertain the physiology of these enigmatic bacteria.

In addition to comammox *Nitrospira*, also two canonical *Nitrospira* MAGs were retrieved from the Breehei DWTP metagenomes, indicating that these nitrite oxidizers can interact with canonical ammonia oxidizers as well as with complete nitrifiers. Moreover, we identified MAGs affiliated with the Verrucomicrobiota and Planctomycetota phyla (Fig. 3) that contained NXR-like sequences with similarity to the enzymes of known NOBs. In phylogenetic analyses of NxrA, they clustered with *Nitrospira*, *Nitrospina,* and anammox bacteria (Figure 5AB). The apparent periplasmic orientation of the NxrA and the lack of transmembrane helices in the NxrC subunit suggests that the NXR of these putative NOBs may be soluble, as has also been proposed for *Nitrosopina gracilis* (Lücker et al., 2013) and “Ct. Nitrotoga fabula” (Kitzinger et al., 2018). However, further studies are needed to analyze the potential nitrite-oxidizing capacity of these bacteria. Although in a very low abundances, *Nitrobacter* species are also detected in some RSFs (Tatari et al., 2017), which we however did not observe in DWTP Breehei.

### 4.3. Methane and one-carbon metabolism

In DWTPs, methane stripping via aeration is preferred over bacterial methane oxidation since microbial activity and growth cause accumulation of extracellular polymeric substances that lead to clogging of the biofilter material (Streese and Stegmann, 2003). However, both WB and P-RSF samples showed high methane-oxidizing capacity (Figure S1). Especially the wall biofilm might counteract methane blowout by oxidizing methane before it leaves the filter and thus reduce methane emissions to the atmosphere. In the global carbon cycle, methane-oxidizing microorganisms play a significant role (Cicerone and Oremland, 1988) as they represent the only known biological methane sink (Aronson et al., 2013). These organisms are able to grow with methane as sole carbon and energy source. The first step of methane oxidation, methane activation and conversion to methanol, is catalyzed either by soluble (sSMO) or particulate (pMMO) methane monoxygenases (Tavormina et al., 2011; Troitsenko and Murrell, 2008). Especially the sSMO is not universal to methanotrophs and only certain phylogenetic groups are known to encode this methane monoxygenase type (Verbeke et al., 2019), usually together with pMMO.

Several studies have shown that the majority of the methane-oxidizing bacteria colonizing the granular material of RSFs are affiliated with the gammaproteobacterial Methylococcaceae family (Albers et al., 2015; Gülay et al., 2016; Palomo et al., 2016). Recently, this family was reclassified and split into the Methylo- monadaceae, Methylococcaceae, and Methylomonadaceae (Parks et al., 2017). Especially members of the Methylo monadaceae have been found in many natural and engineered systems (Flynn et al., 2016; Hoefman et al., 2014; Kalyuzhnaya et al., 2015; Kits et al., 2013; Ogiso et al., 2012; Oswald et al., 2017; Parks et al., 2017; Svenning et al., 2011). In this study, all recovered *Methylo monadaceae* MAGs contained pMMO, and some of them additionally sSMO (Fig. 3, Table S2). In addition, several methanotrophic MAGs also contained the highly divergent pXMO enzyme complex encoded by the *pxmABC* gene cluster (Tavormina et al., 2011). Previous studies have shown that pXMO is involved in methane oxidation under hypoxic denitrifying conditions in *Methylo monadaceae* strains (Kits et al., 2015a, 2015b), but its exact function in the sand filters remains to be determined.

Methane oxidation results in the production of various C₁ intermediates including methanol and formate, which can be used as substrates by methylotrophic bacteria. In DWTP Breehei, the P-RSF samples were dominated by members of the *Methylphilaceae* family, which was assumed to accommodate only methylotrophic bacteria incapable of growth on methane. Surprisingly, one of the *Methylphilaceae* MAGs (P-RSF-NP-02) harbored a complete operon encoding a sSMO (Fig. 3, Table S2), indicating a methanotrophic potential. This finding is substantiated by an earlier study, which demonstrated that a member of methylotrophic genus *Methylococci* can become methanotrophic by acquiring sSMO (Vekeman et al., 2016). The phylogenomic analysis and low ANI value to other members of this family (77.26%) suggest that P-RSF-NP-02 represents a novel species within the *Methylphilaceae* (Fig. 7B) and the extremely high coverage of this MAG indicates a major role in methane removal in the DWTP Breehei. The high abundance of this potential new novel methane-oxidizing bacterium in the P-RSF will be facilitated by the high iron content of the influent water, as the sSMO contains a diiron cluster in the active site (Jasiewski and Que, 2018; Wallar and Lipscomb, 1996).
5. Conclusion

- Metagenomic analyses enabled us to identify key microbial populations involved in the removal of ammonium and methane. While we did not analyze the influence of physicochemical parameters like temperature, pH, or organic carbon content, we found the location within the DWTP Breeheoi to be a major influence factor shaping the microbial community.

- Clade A comammox *Nitrospira* dominated the nitrifying microbial community in P-RSF, while clade B was most abundant in S-RSF, where ammonium concentrations are the lowest, and in the biofilm (WB), which is a predicted niche for comammox bacteria.

- The methanotrophic community was dominated by *Methylphlobacillus* member, which might be highlighted by a high iron concentration in the anaerobic sludge.

Author contributions

LP and HoDc collected and processed samples. LP and HK analyzed and interpreted the data. JF and GC contributed to bioinformatics analyses. TVA and GC performed illumina and Nanopore sequencing. SL, HoDc, MJ and MAJvK were involved in project discussion and data interpretation. SL conceived the research project. LP, HK, and SL wrote the manuscript with input from all the authors.

Data availability

The genome sequences of the 56 MAGs recovered in this study and the raw sequencing data have been deposited in GenBank under BioProject number PRJNA6622654. Nucleotide and amino acid sequences of amoA/amoA, mmoxA and nra/narG genes extracted from the full co-assembly and the recovered MAGs are supplied as supplementary files (Additional data files 1–3, respectively).

Declaration of Competing Interest

The authors declare that they have no competing interests.

Acknowledgments

We would like to thank Weren de Veet, Geert Gielen and Kay Bouts for providing necessary information about DWTP Breeheoi performance and sampling assistance, and Linnea Kop for fruitful discussions. Financial support was provided by the Netherlands Organization for Scientific Research (NWO Talent Programme grants VI.Veni.192.086, 016.Veni.192.062 and 016.Vidi.189.050, and Grantation grants 024.002.001 [NESSC] and 024.002.002 [SIAM]) and the European Research Council (ERC Advanced Grants 339880 [Ecosystem] and 690731 [VOLCANO]).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2020.1162888.

References

Albers, C.N., Elleegaard-Jensen, L., Harder, C.B., Rosendahl, S., Knudsen, B.E., Ekelund, F., Aamand, J., 2015. Groundwater Chemistry Determines the Prokaryotic Community Structure of Waterways Sands Filters. Environ. Sci. Technol. 49 (2), 839–846.

Armenteros, J.J.A., Tsigros, K.D., Sondere, C.K., Petersen, T.N., Wintoner, O., Brunak, S., von Heijne, G., Nielsen, H., 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. 37 (4), 420–423.
Sedlak, D.L.; von Gunten, U. 2011. The Chlorine Dilemma. Science 331 (6013), 42–43.

Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30 (14), 2068–2069.

Sharma, S.K.; Petrushevski, B.; Schippers, J.C. 2005. Biological iron removal from groundwater: a review. J. Water Supply Res. Technol. AQUA 54 (4), 239–247.

Shi, P.; Jia, S.Y.; Zhang, X.X.; Zhang, T.; Cheng, S.P.; Li, A.M. 2013. Metagenomic insights into chlorination effects on microbial antibiotic resistance in drinking wa-
ter. Water Res. 47 (1), 111–120.

Shi, X.Z.; Hu, H.W.; Wang, J.Q.; He, J.Z.; Zheng, C.Y.; Wan, X.H.; Huang, Z.Q. 2018. Niche separation of comammox Nitrospira and canonical ammonia oxidizers in an acidic subtropical forest soil under long-term nitrogen deposition. Soil Biol. Biochem. 126, 114–122.

Simon, J.; Klitz, M.G. 2013. Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. Biochim. Biophys. Acta Bioenerg. 1827 (2), 114–135.

Stackebrandt, E.; Ebers, J. 2006. Taxonomic parameters revisited: tarnished gold standards. Microbiology Today 33 (4), 152–155.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-
analysis of large phylogenies. Bioinformatics 30 (9), 1312–1313.

Stocker, T.F.; Qin, D.; Plattner, G.K.; Tignor, M.; Allen, S.K.; Boschung, J.; Nauels, A.; Migeon, P.M.; Box, V.; Xia, Y. 2013. IPCC: climate change 2013; the Physi-

cal Science Basis. In: Working Group I Contribution to the IPCC 5th Assessment Report. IPCC. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, p. 1353.

Streese, J.; Stegmann, R. 2003. Microbial oxidation of methane from old landfill in biofilters. Waste Manage. 23 (7), 573–580.

Svensson, M.M.; Hedestig, A.G.; Wartenaen, I.; Stein, L.Y.; Klitz, M.G.; Kalyuzh-
naya, M.G.; Spang, A.; Bringel, F.; Vuilleumier, S.; Lajus, A.; Medigue, C.; Bruse, D.C.; Cheng, J.F.; Goodwin, L.; Ivanova, N.; Han, J.; Han, C.S.; Hausier, L.; Held, B.; Land, M.L.; Lapidus, A.; Lucas, S.; Nolan, M.; Pitluck, S.; Wooley, T. 2011. Genome Sequence of the Arctic Methanotroph Methylobacter tundripaludum SVM9. J. Bacteriol. 193 (22), 6418–6419.

Tatari, K.; Musovic, S.; Guay, A.; Deschesne, A.; Albrechtens, H.J.; Smets, B.F. 2017. Density and distribution of nitrifying guilds in rapid sand filters for drinking water production: dominance of Nitrospira spp. Water Res. 127, 239–248.

Tavormina, P.L.; Orphan, V.J.; Kalyuzhnaya, M.G.; Jetten, M.S.M.; Klitz, M.G. 2011. A novel family of functional operons encoding methane/ammonia monoox-
ynase-related proteins in gammaproteobacterial methanotrophs. Environ. Micro-

b. Rep. 3 (1), 91–100.

Trifinopoulos, J.; Nguyen, L.T.; von Haeseler, A.; Minh, B.Q. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44 (W1), W232–W235.

Trottenko, Y.A.; Murrell, J.C. 2008. Metabolic aspects of aerobic obligate methane-

otrophy. Adv. Appl. Microbiol. 63, 183–229.

Trussell, R.R.; Hand, D.W.; Tchobanoglous, G.; Howe, K.J.; Crittenden, J.C. 2012. Principles of Water Treatment. John Wiley & Sons Inc, Hoboken, New Jersey.

van Kessel, M.; Speth, D.; Albertsson, M.; Nielsen, P.H. 2010. Op den Camp, H.J.M.; Kartal, B.; Jetten, M.S.M.; Lücke, S. 2015. Complete nitrification by a single microorganism. Nature 528, 555–559.

van der Wielen, P.W.J.; Voost, S.; van der Kooij, D. 2009. Ammonia-Oxidizing Bact-
teria and Archaea in Groundwater Treatment and Drinking Water Distribution Systems. Appl. Environ. Microbiol. 75 (14), 4687–4695.

Vaske, R.; Sozic, I.; Nagarajan, N.; Sikic, M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 27 (5), 737–746.

Vekeman, B.; Berkholz, F.M.; Cremer, G.; de Vos, P.; Vandamme, P.; Boon, N. 2006. Op den Camp, H.J.M.; Heylen, K. 2016. New Methylococaceae diversity from North

Sea sediments includes methanotroph containing solely the soluble methane

monooxygenase. Environ. Microbiol. 18 (12), 4523–4536.

Verbeke, T.J.; Dedvy, S.N.; Dimond, P.F. 2018. Methanotrophy in Acidic Soils, In:

cluding Northern Peatlands. In: McGinity, T. (Ed.), Microbial Communities Util-
izing Hydrocarbons and Lipids: Members, Metagenomes and Ecophysiology. Handbook of Hydrocarbon and Lipid Microbiology. Springer, Cham. doi:10.1007/978-3-319-14783-5_6.

Wagner, F.B.; Nielsen, P.B.; Bøe-Hansen, R.; Albrechtens, H.J. 2016. Copper deficiency can limit nitrification in biological rapid sand filters for drinking water produc-

ing. Water Res. 95, 280–288.

Wallar, B.J.; Lipscomb, J.D. 1996. Oxygen activation by enzymes containing binu-

clear non-heme iron clusters. Chem. Rev. 96 (7), 2625–2657.

Wang, Y.; Ma, L.; Mao, Y.; Jiang, X.; Xia, Y.; Yu, K.; Li, B.; Zhang, T. 2017. Comammox in drinking water systems. Water Res. 116, 332–341.

Wang, Z.H.; Cao, Y.Q.; Zhu-Barker, X.; Nicol, G.W.; Wright, A.L.; Jia, Z.J.; Jiang, X.J. 2019. Comammox Nitrospira clade B contributes to nitrification in soil. Soil Biol. Biochem. 135, 392–395.

Whitman, W.B.; Oren, A.; Chuvashina, M.; da Costa, M.S.; Garrison, G.M.; Rainey, F.A.; Rossello-Mora, R.; Schink, B.; Suttle, I.; Trümpi, J.; Ventura, S. 2018. Pro-

dosal of the suffix -ota to denote phyla. Addendum to ‘Proposal to include the rank of phylum in the International Code of Nomenclature of Prokaryotes’. Int. J. Syst. Evol. Microbiol. 68 (3), 967–969.

WHO 2003. Ammonia in drinking-water. Background document for Preparation of WHO Guidelines for Drinking-Water Quality. World Health Organization [WHO/SEDE/WHSH03/04/01], Geneva.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. SpringerSpringer,

Heidelberg, Germany.

Wilczak, A.; Jarczega, J.G.; Marcincz, J.P.; Odell, L.H.; Kirmeyer, G.J.; Wolfe, R.L. 1996. Occurrence of nitrification in chlorinated distribution systems. J. Am. Water Works Assoc. 88 (7), 74–85.

Yang, J.X.; Ma, J.; Song, D.; Zhi, X.D.; Kong, X.J. 2016. Impact of preozonation on the bioactivity and biodiversity of subsequent biofilters under low temperature conditions: A pilot study. Front. Environ. Sci. Eng. 10 (4).

Zhang, H.H.; Chang, F.Y.; Shi, P.; Ye, L.; Zhou, Q.; Pan, Y.; Li, A.M. 2019. Antibiotic Resistome Alteration by Different Disinfection Strategies in a Full-Scale Drinking Water Treatment Plant Deciphered by Metagenomic Assembly. Environ. Sci. Technol. 53 (4), 2141–2150.

Zhou, J.Z.; Bruns, M.A.; Tiedje, J.M. 1996. DNA recovery from soils of diverse composition. Appl. Environ. Microbiol. 62 (2), 316–322.