Ecological patterns, diversity and core taxa of microbial communities in groundwater-fed rapid gravity filters

Gülay, Arda; Musovic, Sanin; Albrechtsen, Hans-Jørgen; Abu Al-Soud, Waleed; Sørensen, Søren Johannes; Smets, Barth F.

Published in:
The ISME Journal

DOI:
10.1038/ismej.2016.16

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
Gülay, A., Musovic, S., Albrechtsen, H.-J., Abu Al-Soud, W., Sørensen, S. J., & Smets, B. F. (2016). Ecological patterns, diversity and core taxa of microbial communities in groundwater-fed rapid gravity filters. The ISME Journal, 10(9), 2209-2222. https://doi.org/10.1038/ismej.2016.16

Download date: 06. Aug. 2023
Ecological patterns, diversity and core taxa of microbial communities in groundwater-fed rapid gravity filters

Arda Gülay1, Sanin Musovic1, Hans-Jørgen Albrechtsen1, Waleed Abu Al-Soud2, Søren J Sørensen2 and Barth F Smets1

1Department of Environmental Engineering, Technical University of Denmark, Kongens Lyngby, Denmark and 2Department of Biology, University of Copenhagen, Copenhagen, Denmark

Here, we document microbial communities in rapid gravity filtration units, specifically serial rapid sand filters (RSFs), termed prefilters (PFs) and after-filters (AFs), fed with anoxic groundwaters low in organic carbon to prepare potable waters. A comprehensive 16S rRNA-based amplicon sequencing survey revealed a core RSF microbiome comprising few bacterial taxa (29–30 genera) dominated by Nitrospirae, Proteobacteria and Acidobacteria, with a strikingly high abundance (75–87 ± 18%) across five examined waterworks in Denmark. Lineages within the Nitrospira genus consistently comprised the second most and most abundant fraction in PFs (27 ± 23%) and AFs (45.2 ± 23%), respectively, and were far more abundant than typical proteobacterial ammonium-oxidizing bacteria, suggesting a physiology beyond nitrite oxidation for Nitrospira. Within the core taxa, sequences closely related to types with ability to oxidize ammonium, nitrite, iron, manganese and methane as primary growth substrate were identified and dominated in both PFs (73.6 ± 6%) and AFs (61.4 ± 21%), suggesting their functional importance. Surprisingly, operational taxonomic unit richness correlated strongly and positively with sampling location in the drinking water treatment plant (from PFs to AFs), and a weaker negative correlation held for evenness. Significant spatial heterogeneity in microbial community composition was detected in both PFs and AFs, and was higher in the AFs. This is the first comprehensive documentation of microbial community diversity in RSFs treating oligotrophic groundwaters. We have identified patterns of local spatial heterogeneity and dispersal, documented surprising energy–diversity relationships, observed a large and diverse Nitrospira fraction and established a core RSF microbiome.

The ISME Journal (2016) 10, 2209–2222; doi:10.1038/ismej.2016.16; published online 8 March 2016

Introduction

Microbial ecology seeks to find out what forces shape and maintain the composition and structure of microbial communities; elucidate the origin of community and guild diversity and their link to system function, stability and resilience; and identify the interactions that stabilize or constrain microbial communities; among others. Answering those questions requires access to microbial communities that can be observed, sampled and described at relevant spatiotemporal scales, and with sufficient replication, while the system’s environmental conditions and performance can be measured. Technical systems—where open microbial communities are subject to some degree of management—might provide useful alternatives to field systems—without suffering from the simplicity of laboratory-based model systems. Here, we introduce the microbial communities that populate rapid sand filters (RSFs), which are constantly fed with anoxic groundwater (GW) after vigorous aeration, as model technical systems, and present our initial analysis on their composition and ecological patterns.

Rapid gravity filtration is a conventional treatment process whereby GW is intentionally passed through a porous medium bed, often consisting of sand (then called rapid sand filtration), to produce high-quality drinking water. Rapid gravity filters receive continuous inputs from aquifers over multiple years to decades and this input may vary depending on the temporal and spatial dynamics of the aquifer (Griebler and Lueders, 2009). Conditions in source aquifers are similar in that they lack the photosynthetic activity, are at low temperature, carry no dissolved oxygen and typically have low contents of organic carbon, yet they can differ in their content of inorganic (e.g., NH₄⁺, Mn²⁺, Fe²⁺, H₂S) as well as some organic electron donors (mainly CH₄) (Danielpol et al., 2000), all potential energy sources for microbial growth.
Studies on the diversity, abundance and distribution of microorganisms in GW-fed RSFs have only recently been reported. Using 16S rRNA gene cloning, van der Wielen and van der Kooij (2009) and White et al. (2012) identified *Nitrosomonas oligotropha*, *Nitrosomonas marina* and *Candidatus Nitrosopumilus maritimus* as the dominant ammonia-oxidizing prokaryotes in three GW treatment plants. Pinto et al. (2012) and Albers et al. (2015), using 16S rRNA pyrosequencing, identified *Proteobacteria*, *Acidobacteria*, *Actinobacteria*, *Nitrospiridae* and *Plantomycetes* as the major phyla in the filtration units. Using a 16S rRNA-targeted DGGE approach, de Vet et al. (2009) identified a shift from iron-oxidizing bacterial dominance in the subsurface-aerated GW to ammonia-oxidizing archaea in above-ground RSF. In a previous study, using 16S rRNA pyrosequencing, we identified *Nitrosomonas europaea*, *Candidatus Nitrosoarchaeum* and *Nitrospira moscoviensis* as dominant nitrifiers in a GW-based full-scale RSF (Gülay et al., 2014).

Reports on RSF microbial communities have typically involved limited sampling sites with insufficient replication within and across RSFs and across drinking water treatment plants (DWTPs) to allow inference of distribution patterns, identify the degree of filter-specific versus shared taxa across DWTPs or assess spatial variation of community structure within DWTPs.

We contend that further knowledge about the microbial community structure of RSFs and its core members is important, and RSF communities can provide suitable model systems to answer central microbial ecological questions. Core taxa have been identified in anaerobic digesters (Riviére et al., 2009), marine environments (Pommier et al., 2007) and activated sludge systems (Zhang et al., 2012; Saunders et al., 2015). Distribution patterns of shared taxa can reveal to what extent environmental conditions vis-à-vis dispersal drive microbial community structure across habitats (Shade et al., 2012) and help identify the microbial taxa that govern central system functions. In addition, the effects of dispersal limitation (Bell, 2010), environmental heterogeneity (Ramette and Tiedje, 2007), energy input (Bienhold et al., 2011) and system design and operation (Sundberg et al., 2013) on microbial community structure are central to many environments, and RSFs might be ideally suited for their investigation.

Here, we extensively characterize the microbial communities at five different DWTPs. Both the feed GWs and the serial filtration units (termed prefilters (PFs) and after-filters (AFs)) were analyzed via 16S rRNA amplicon sequencing. Replicate samples were taken within individual filters and between filters. We first identified and quantified taxa shared among DWTPs. We then evaluated to what extent energy availability could explain bacterial richness and evenness patterns within DWTPs. We also explored distance-decay relationship between aquifer microbial communities and the extent of dispersal within a DWTP. Finally, we examined the extent of local (within filter) spatial heterogeneity in microbial community structure at all RSFs as initiated before (Gülay and Smets, 2015).

### Materials and methods

#### Study sites

Filter materials and GW samples were obtained from five DWTPs in Denmark (Supplementary Figure S1): first line of Langerød (DWTP1, 55°41′45″N, 11°40′44″E); second line of Langerød (DWTP2); Sjælsø-1 (DWTP3, 55°52′25″N, 12°28′36″E); Sjælsø-2 (DWTP4, 55°52′20″N, 12°28′32″E); Islevbro (DWTP5, 55°42′14″N, 12°27′21″E). Study sites are detailed in Supplementary Information.

#### Sampling procedures and physicochemical characteristics of GW and discharge water

Typical values of the dominant chemical constituents in the GW and performance at the different DWTPs are in Table 1. Water quality data (between 2009 and 2011) of raw and discharge water were obtained from the JUPITER database (Hansen and Pjetursson, 2011) and from direct contact with the water utilities, respectively. GW samples were collected from the main influent lines under anoxic conditions before the aeration unit. Microbial mass was then harvested from 250 ml GW using a sterile 0.2 μm pore size (Millipore, Billerica, MA, USA) membrane filter and stored at –20 °C. Filter samples were collected by inserting a 60-cm-long core sampler at a single PF and two separate AF units at each DWTP 6 days AF backwash (Supplementary Figure S2). Sampling procedures are detailed in the Supplementary Information.

#### DNA extractions, PCR amplification and pyrosequencing

DNA was extracted, amplified and subject to 16S rRNA amplicon 454 pyrosequencing, essentially following the protocol by Sundberg et al. (2013), as described in the Supplementary Information.

#### Real-time qPCR

Quantification of total bacteria (Eubacteria), the sum of Nitrosomonas plus Nitrospira, and Nitrospira was performed by 16S rRNA gene targeted quantitative PCR (qPCR) on replicates of DNA extracts and the results are described elsewhere (Tatari et al., in preparation). The qPCR primers, PCR conditions and procedures used are described in the Supplementary Information. Results, as reported by Tatari et al. (in preparation), are used here to calculate relative and absolute abundance values.
The ISME Journal

Table 1 Influent conditions and performance of the studied drinking water treatment plants

| Compound  | Effluent guidelines | DWTP1 ± 2 | DWTP3 | DWTP4 | DWTP5 |
|-----------|---------------------|-----------|-------|-------|-------|
| Temperature (°C) | — | 9.1 ± 0.4 | 9.5 ± 0.3 | 8.8 ± 0.3 | 9.3 ± 0.4 |
| pH | 7.0–8.5 | 7.4 ± 0.1 | N/A | 7.5 ± 0.1 | 7.4 ± 0.3 |
| Dissolved oxygen (mg O2 l⁻¹) | 5 | 0.26 ± 0.11 | 0.65 ± 0.27 | 0.87 ± 0.44 | 1.09 ± 0.59 |
| Ammonium (mg NH₄l⁻¹) | 0.05 | 1.06 ± 0.18 | 0.51 ± 0.21 | 1.08 ± 0.44 | 0.41 ± 0.2 |
| Iron (mg Fe l⁻¹) | 0.1 | 3.0 ± 0.2 | 2.9 ± 1.7 | 1.9 ± 1.1 | 1.8 ± 1.1 |
| Manganese (mg Mn l⁻¹) | 0.02 | 0.115 ± 0.010 | 0.14 ± 0.09 | 0.07 ± 0.06 | 0.06 ± 0.02 |
| Hydrogen sulfide (mg H₂S l⁻¹) | 0.05 | 0.02 ± 0.00 | 0.03 ± 0.02 | 0.12 ± 0.27 | 0.02 ± 0.00 |
| Methane (mg CH₄ l⁻¹) | 0.01 | 0.06 ± 0.03 | 0.12 ± 0.37 | 3.73 ± 2.67 | 0.02 ± 0.03 |
| Phosphorus (mg total P l⁻¹) | 0.15 | 0.269 ± 0.051 | N/A | N/A | 0.025 ± 0.01 |
| Bicarbonate (mg HCO₃ l⁻¹) | — | 361 ± 6 | N/A | N/A | 346 ± 48 |
| Chloride (mg Cl l⁻¹) | 250 | 55 ± 15 | 51 ± 36 | 41 ± 15 | 64 ± 32 |
| Sulfate (mg SO₄ l⁻¹) | 250 | 7 ± 3 | 36 ± 22 | 12 ± 17 | 102 ± 28 |
| NVOC (mg C l⁻¹) | 4 | 2.6 ± 0.3 | 2.7 ± 0.4 | 2.3 ± 0.3 | 2.3 ± 0.3 |
| Treatment plant performance | — | Problems with incomplete ammonium removal | None of the above parameters exceeded the effluent guideline from 2009 until 2011 | None of the above parameters exceeded the effluent guideline from 2009 until 2011 |

Abbreviations: DWTP, drinking water treatment plant; N/A, no data available; NVOC, nonvolatile organic carbon.
Raw water values for DWTP3 and DWTP1–2 are weighted averages ± s.d. (weighting based on raw water abstraction data). Raw water values for DWTP3 and DWTP4 are average values ± s.d.

Bioinformatic and statistical analysis

Raw sequence data were quality checked (denoised) with Ampliconnoise (Quince et al., 2011) and all analyses were performed using the QIIME 1.5.0 software (Caporaso et al., 2010). Subsequent bioinformatics and statistical analyses are described in the Supplementary Information, and β-diversity significance testing was performed as described earlier (Gülay and Smets, 2015).

Nucleic acid sequences

Raw SFF files were deposited into the Sequence Read Archive at GenBank under the study accession number SRP045492.

Results

Overall measures of diversity

A total of 8,578,355 (74.7% of total) amplicon sequences (average length 421 nucleotides) passed all quality checks, including denoising and chimera removal, and were distributed across 33,019 operational taxonomic units (OTUs), defined at 97% sequence identity, hereafter called OTU₀.₀₃. We rarefied to 4,000 sequences per sample for further diversity and taxonomic comparisons.

While rarefaction analysis suggests some undersampling of the metacommunity at all DWTPs based on individual or merged samples (Supplementary Figures S3 and S4), the large degree of replication provided many subsamples of the metacommunity. Comparison of microbial diversity in RSFs with other environments showed a high degree of richness in RSFs, notwithstanding the much lower input of nutrients (Supplementary Table S2).

The Shannon index of the GW community was twofold higher compared with that of the PF and AF communities; consistently, the Gini coefficient was lowest in the GW communities. Hence, the GW community had the most even distribution of OTU₀.₀₃s. Of all identified OTU₀.₀₃s, 99% (32,627) could be taxonomically assigned using NCBI BLASTN analyses (Altschul et al., 1997) at a minimum identity of 90%. Thousand seventy-five taxa were identified at the genus level from 21,941 OTU₀.₀₃s after sample size normalization. The communities were dominated by bacteria comprising 99.9%, 99.5% and 99.7% of all sequences in GW, PF and AF, respectively (Supplementary Table S3). At the phylum level, a large fraction of the sequences in PFs and AFs (89.56 ± 19% and 87.62 ± 15%, respectively) were Nitrospirae, Proteobacteria and Acidobacteria (Figure 1). Proteobacterial OTU₀.₀₃s in
### Table 2

| Water Works | Unit | Zone | Number of Sequences<sup>a</sup> | Singleton OTU<sub>b</sub> | Chao1 | ACE | Gini<sub>c</sub> | Diversity |
|-------------|------|------|----------------------------------|--------------------------|-------|-----|----------------|-----------|
| DWTP1       | GW   | Top  | 6758                             | 372                      | 192   | 1005| 890           | 0.841     | 8.5 ± 0.02|
|             | PF   | Top  | 12,050 ± 2,396                   | 322 ± 4                  | 175 ± 3| 546 ± 81| 641 ± 99      | 0.890 ± 0.01 | 4.1 ± 0.18|
|             |      | Depth| 12,414 ± 3,826                   | 361 ± 10                 | 192 ± 34| 604 ± 130| 742 ± 150      | 0.881 ± 0.01 | 4.2 ± 0.24|
|             | AF   | Top  | 13,908 ± 7,873                   | 439 ± 66                 | 257 ± 56| 779 ± 130| 909 ± 161      | 0.862 ± 0.01 | 4.8 ± 0.17|
|             |      | Depth| 14,361 ± 1,988                   | 497 ± 69                 | 296 ± 47| 959 ± 155| 1100 ± 178     | 0.855 ± 0.01 | 4.8 ± 0.40|
| DWTP2       | GW   | Top  | 6392                             | 1105                     | 772   | 3046| 3352          | 0.683     | 6.1 ± 0.03|
|             | PF   | Top  | 7738 ± 553                       | 393 ± 47                 | 284 ± 31| 1142 ± 182| 1337 ± 182     | 0.895 ± 0.01 | 4.3 ± 0.20|
|             |      | Depth| 7113 ± 622                       | 406 ± 30                 | 286 ± 26| 1406 ± 59| 1410 ± 97      | 0.888 ± 0.01 | 4.5 ± 0.05|
|             | AF   | Top  | 9686 ± 3,891                     | 512 ± 45                 | 340 ± 49| 1361 ± 362| 1525 ± 330     | 0.862 ± 0.01 | 4.6 ± 0.09|
|             |      | Depth| 11,757 ± 4,330                   | 480 ± 62                 | 301 ± 71| 1275 ± 473| 1371 ± 505     | 0.868 ± 0.01 | 4.5 ± 0.02|
| DWTP3       | GW   | Top  | 18746                            | 863                      | 527   | 1695| 1895          | 0.775     | 6.8 ± 0.04|
|             | PF   | Top  | 6803 ± 1,332                     | 470 ± 69                 | 343 ± 55| 1639 ± 302| 1719 ± 336     | 0.929 ± 0.01 | 3.1 ± 0.60|
|             |      | Depth| 6968 ± 751                       | 564 ± 50                 | 427 ± 58| 1893 ± 233| 2241 ± 401     | 0.879 ± 0.04 | 3.3 ± 0.43|
|             | AF   | Top  | 6406 ± 1,228                     | 802 ± 81                 | 589 ± 79| 2694 ± 494| 3008 ± 520     | 0.721 ± 0.03 | 7.6 ± 0.01|
|             |      | Depth| 6803 ± 1,332                     | 784 ± 105                | 590 ± 114| 2580 ± 591| 2973 ± 706     | 0.716 ± 0.03 | 7.5 ± 0.41|
| DWTP4       | GW   | Top  | 16,480                           | 1169                     | 676   | 2170| 2373          | 0.667     | 8.6 ± 0.03|
|             | PF   | Top  | 6397 ± 930                       | 239 ± 45                 | 161 ± 39| 750 ± 202| 770 ± 296      | 0.879 ± 0.01 | 4.0 ± 0.37|
|             |      | Depth| 7498 ± 888                       | 357 ± 42                 | 238 ± 34| 1222 ± 165| 1311 ± 236     | 0.852 ± 0.01 | 4.6 ± 0.26|
|             | AF   | Top  | 6880 ± 952                       | 1106 ± 192               | 850 ± 184| 4368 ± 1194| 4932 ± 1526   | 0.784 ± 0.02 | 6.4 ± 0.08|
|             |      | Depth| 6390 ± 782                       | 1104 ± 176               | 844 ± 164| 4599 ± 1715| 5013 ± 1689    | 0.792 ± 0.01 | 6.3 ± 0.47|
| DWTP5       | GW   | Top  | 16,648                           | 962                      | 593   | 1808| 2183          | 0.735     | 7.5 ± 0.03|
|             | AF   | Top  | 8084 ± 740                       | 343 ± 68                 | 241 ± 45| 1131 ± 317| 1233 ± 354     | 0.909 ± 0.01 | 3.4 ± 0.36|
|             |      | Depth| 7650 ± 10,26                     | 562 ± 161                | 422 ± 142| 1913 ± 742| 2136 ± 853     | 0.859 ± 0.04 | 4.6 ± 1.02|

Abbreviations: ACE, abundance-based coverage estimator; AF, after-filter; DWTP, drinking water treatment plant; G<sub>c</sub>, Gini coefficient; GW, groundwater; H, shannon index; OTU, operational taxonomic unit; PF, prefilter.

<sup>a</sup> All sample sizes were equalized at 4000 sequences. Standard deviations (±) were obtained from biological replicates; standard deviations in the top and bottom layers were assessed from triplicate samples from the same filter. GW samples were taken as individual samples.

<sup>b</sup> Numbers of sequences in PF and AF rows show the total amount of high-quality sequences that were obtained from 10 ng of DNA extracted from 0.5 g of drained wet sand grains. Mean values were calculated by grouping the samples according to zones, units and waterworks.

<sup>c</sup> OTU calculations were carried out at 97% phylogenetic similarity.

---

The PFs and AFs were distributed across the gamma (19.8 ± 15%), alpha (13.9 ± 5%), beta (7.4 ± 5%) and delta (4.7 ± 7%) subphyla.

In PFs and AFs, absolute bacterial abundances, expressed in gene copies per gram of filter material, based on broad and clade-specific 16S rRNA-targeted qPCR ranged from 2 to 12 and 1 to 4 × 10<sup>9</sup> total bacteria, 6 to 23 and 0.4 to 15 × 10<sup>7</sup> AOB (Nitrosomonas plus Nitrosospira), 8 to 370 and 3 to 19 × 10<sup>4</sup> Nitrobacter and 2 to 24 and 1 to 9 × 10<sup>9</sup> Nitrosospira, respectively, across the examined DWTPs (Tatari et al., in preparation). Comparing qPCR results on individual samples confirmed the high relative abundance of Nitrosospira in all samples (Figure 2a). The qPCR estimated abundances are somewhat lower compared with those estimated by amplicon sequencing, but overrepresentation of abundant taxa when using universal primers in PCR-based surveys is well documented (Gonzalez et al., 2012). Nevertheless, even qPCR estimates reveal an extremely high Nitrosospira (on average 17 ± 11.6%) and much lower AOB abundance (2.26 ± 2.79%) consistent with amplicon sequencing estimates (Figure 2b).

The microbial communities in the PFs were dominated (30.3 ± 24%) by sequences of an uncultured genus of the Methylococcales (γ-proteobacterial methane oxidizers) and this taxon decreased in abundance in the AF communities (5.7 ± 3%). The most abundant bacterial lineage in the AFs (45.2 ± 23%) and the second most abundant in the PFs (27 ± 23%) was the genus Nitrospira (phylum Nitrospirae), typically known as nitrite oxidizer (Pester et al., 2014).

**Core taxa among water works**

We speculated that similarities in influent, design and operational parameters of the DWTPs would lead to dominant DWTP-shared RSF microbial communities. Hence, shared taxa in the microbial communities of the GW, the PFs and the AFs at all DWTPs were determined. Shared taxa were classified at the phylum, class, order and genus level, and taxa (at the genus level) present at all DWTPs at more or >1% mean relative abundance were defined as dominant or rare core taxa (Gobet et al., 2010), respectively (Figure 3).
At the genus level, the GW comprised 73 core taxa (9.5% of all core genera) in 14 phyla accounting for 68.9 ± 13% of the total sequence abundance (Figure 3a). In the PF microbial communities, 30 core taxa (4% of all core genera) were identified, comprising 87 ± 5% of the total sequence abundance (Figure 3a). Most of the PF core taxa (72% of core genera) were also in the GW core (Figure 3b), although the habitats of anoxic GWs and highly oxic PFs are different. Twenty-nine taxa (3% of all core genera) constituted the core of AF microbial communities, representing 75 ± 18% of the total sequence abundance (Figure 3a). Six of the rare core taxa in the PF were dominant core taxa in the AF, including Nitrosomonadales, OM190, Chloracidobacterium, Xanthomonicales, Anaerolineales and Da023. Both PF and AF community cores were dominated by Caulobacteriales, Nitrosirales, Methylococcales, Rhizobiales and Myxococcales.

Although few genera constituted the core in both PF and AF communities, the number of OTUs per core genus varied significantly. The number of OTUs per taxon exponentially related to its abundance (i.e., sequence number) in both PF...
(\(r^2 = 0.75; P = 4.21 \times 10^{-08}\)) and AF communities (\(r^2 = 0.55; P = 1.03 \times 10^{-04}\) with Da023, \(r^2 = 0.82; P = 3.39 \times 10^{-08}\) without Da023), except for genus Da023. The sequence abundance and intragenus OTU\(_{0.03}\) richness of Nitrospira, Methylosoma, Hypomicrobium, Rhizobiales (uncultured) and Caulobacterales (uncultured) was, thus, highest of all core taxa. Da023 displays an anomalously high OTU\(_{0.03}\) richness for its total sequence abundance of 8674 across all AF communities (on average abundance 131 ± 69) (Supplementary Figure S6).

**Physiological typing of the core taxa**

Phylogenetic trees were constructed containing sequences from the core taxa and their closest cultured relatives. Assuming shared metabolic pathways among phylogenetically related organisms, this suggests a physiology of the core taxa (Langille et al., 2013). A high number of OTU\(_{0.03}\)s within the core taxa of the PFs (eight taxa) and AFs (seven taxa) were closely related (≥95% sequence similarity) to sequences of types with the ability to oxidize ammonium, nitrite, iron, manganese and methane as primary growth substrate (Figure 4). Furthermore, these OTU\(_{0.03}\)s represent up to 79 ± 2.3% and 66 ± 0.6% of all sequences in the PF and AF core, respectively (Table 3).

In PFs and AFs, 31 OTU\(_{0.03}\) had Nitrosomonas oligotropha, a high-affinity chemolithotrophic ammonia oxidizer (Koops et al., 2006) as closest relative; in addition, Nitrosoccus mobilis, another oligotrophic ammonium oxidizer, was a close phytype to many sequences in PF (25 OTU\(_{0.03}\)s), and Nitrosomonas sp. BF16c52, a Nitrosomonas mobilis strain isolated a nitrifying freshwater system (Burrell et al., 2001) clustered with 73 OTU\(_{0.03}\)s of AF core taxa (Figure 4). Within the nitrite-oxidizing clades, a majority of sequences phylogenetically affiliated with uncultured representatives of the Nitrospira genus (169 OTU\(_{0.03}\)s in the PF core, and 243 in the AF core). In addition, less OTU\(_{0.03}\)s grouped with Cand. Nitrospira defluvii (25) and Nitrospira moscoviensis (80) in the PF and Nitrospira marina (36), Nitrospira calida (105) and Cand. Nitrospira defluvii (1) in the AF.

In PFs, 43 OTU\(_{0.03}\)s of the core genera had Crenothrix polyspora, a methane-oxidizing bacterium
originally isolated from an RSF (Wolfe, 1960), as a closest relative. The presence of such phylotypes indicates methane residuals in the influent water, although aeration units preceding the PFs have methane removal as central objective. *C. polyspora*-related phylotypes were not detected in the AF core; here, abundant taxa were identified with another methane oxidizer *Methylovulum miyakonense* (43) and uncultured representatives of the *Methylovurum* genus as closest relatives. OTU0.03s with iron- and manganese-oxidizing strains as closest phylotypes were mostly detected in the PF core. Thirty-two

---

**Table 3** Abundance of taxa (%) with known physiology in the core taxa at each waterworks

|                | Prefilter |          |          |          | After-filter |          |          |          |          |
|----------------|-----------|----------|----------|----------|-------------|----------|----------|----------|----------|
|                | DWTP1     | DWTP2    | DWTP3    | DWTP4    | DWTP1       | DWTP2    | DWTP3    | DWTP4    | DWTP5    |
| AOB            | 1.55      | 4.00     | 3.50     | 0.09     | 8.79        | 2.54     | 1.00     | 0.12     | 0.11     |
| NOB            | 9.82      | 29.24    | 11.89    | 73.64    | 60.40       | 70.49    | 44.97    | 19.11    | 71.45    |
| MOB            | 54.46     | 39.70    | 62.69    | 4.09     | 2.37        | 4.26     | 9.66     | 5.76     | 4.09     |
| IOB            | 12.77     | 6.72     | 0.73     | 0.23     | 2.37        | 4.26     | 9.66     | 5.76     | 4.09     |
| MnOB           | 0.16      | 0.15     | 0.24     | 0.38     | —           | —        | —        | —        | —        |
| SOB            | 0.70      | 0.53     | 0.60     | 0.20     | —           | —        | —        | —        | —        |
| Total          | 79.46     | 80.34    | 79.65    | 78.62    | 71.55       | 77.29    | 55.63    | 25.00    | 75.65    |

Abbreviations: AOB, ammonium oxidizers; DWTP, drinking water treatment plant; IOB, iron oxidizers; MOB, methane oxidizers; MnOB, manganese oxidizers; NOB, nitrite oxidizers; SOB, sulfide oxidizers.
OTU_{0.03}s had *Pedomicrobium manganicum*, a dimorphic prosthecate and oligotrophic manganese-oxidizing bacterium (Poindexter, 2006), as a closest relative, whereas 15 OTU_{0.03}s clustered with sequences assigned to uncultured *Gallionellaceae* putatively strains with iron-oxidizing capability.

**Relationships of diversity with energy availability**

We investigated whether richness and evenness patterns in the RSFs related to energy availability within DWTPs. Energy sources for chemolithotrophic growth are gradually consumed between the serial filtration units (Lopato et al., 2013; Gülay et al., 2014), and within individual units removal is spatially stratified (Lopato et al., 2011, 2013; Tatari et al., 2013; Lee et al., 2014). Hence, a gradient in energy supply is present along the four sampling points in the DWTPs: from top to bottom of the PF and from top to bottom of the AF. In DWTP2, 3 and 4, decreases in total bacterial abundance were estimated from PF to AF, whereas in DWTP5, where typical AF units were absent, total bacterial density was clearly stratified with filter depth (Tatari et al., in preparation). Abundance of the *Nitrosomonas* plus *Nitrospira* clade, followed the pattern of total bacterial abundance: decreasing from PF to AF units at DWTPs 2, 3, 4, whereas total abundance of the *Nitrospira* clade also decreased at DWTPs 3 and 4, but significantly increased from PF to AF at DWTPs 1 and 2 (Tatari et al., in preparation).

At community level, OTU richness correlated strongly and positively with sampling location in the DWTP (r^2\_adj. = 0.69, P = 0.0001; Figure 5a), whereas a weaker and negative correlation was observed for OTU evenness (Gini coefficients, r^2\_adj. = 0.11, P = 0.003; Figure 5b). This pattern was consistent across DWTPs and at each DWTP: OTU richness in all five DWTPs increased from PF to AF (r^2\_adj. = 0.57 ± 0.2, P < 10^-2; Supplementary Figure S7) and Gini coefficient significantly decreased (r^2\_adj. = 0.54 ± 0.12, P < 10^-2; Supplementary Figure S8).

---

**Figure 5** Ecological patterns in RSFs. Changes in OTU richness (a) and evenness (b) along sampling location. (c) Sequence abundance versus OTU richness of phyla at four sampling locations across DWTPs. Each individual dot represents a phylum at a sampling location. (d) Phylogenetic dissimilarity between *Nitrospira* OTUs across aquifers, PFs and AFs of DWTPs, based on unweighted UniFrac diversity estimates. (e) β-diversity analysis applied to six biological replicate samples from a filter unit based on weighted UniFrac diversity estimates. Corrected β-diversity values of replicate samples (black circles) and estimated meta-community subsamples (red circles) are plotted.
For 4 and 21 core genera (of 30 total) in both PFs and AFs, evenness (0.17 ± 0.04) and richness (0.53 ± 0.24) were related to sampling location along the DWTP, respectively (Supplementary Table S4 and Supplementary Figure S9). Richness decreased with location for genera Akivs43, Crenothrix, Thiobacillus, Pedomicrobium, OD1, Pir and uncultured representatives of Burkholderiales, Nitrosomona- dales, Rhodospirillales; richness increased with location for genera Haliangium, Chloracidobacter- ium, Da023, Nitrospira, Om190 and uncultured representatives of Xanthomonadales, Anaerolinea- les and Acidimicrobiales (Supplementary Figure S9).

Spatial heterogeneity
The microbial community structures at different sampling locations within an individual RSF were all significantly different (p-value exceeded the metacommunity-derived threshold) at all DWTPs (Figure 1e), and the magnitude of dissimilarities were similar across PFs and AFs (Supplementary Figure S11). In PFs, however, communities were on average less spatially heterogeneous than in AFs, both in the horizontal plane and with depth. Community dissimilarity between DWTPs was greater than within DWTPs, and within a DWTP dissimilarities decreased along the treatment train (from GW to AF) (Supplementary Figure S12).

Microbial dispersal within DWTPs and between aquifers
Logically, microbial communities from GWs would seed and influence the taxonomic composition of the PFs communities; similarly, the PFs would seed the AFs. To assess this occurrence, each dominant OTU in a unit was checked for its occurrence in the upflow units. Dominant phylotypes (OTU0.03s) in the downflow units should also occur in the upflow units, if seeding were significant. On average, 75.4 ± 20% of the dominant OTU0.03s in PF communities were detected in the GWs, be it as rare members. For the dominant OTU0.03s in the AFs, 99 ± 2% were found as PF OTU0.03s, of which 49% were dominant and 51% rare (Supplementary Table S5). Clearly, the taxonomic inventory of the community in upflow process units, including the rare members, affects the downflow units.

Second, because deep aquifers are isolated eco- systems (Fredrickson and Onstott, 2001), a decay in taxon similarity is expected with distance. We investigated the degree of phylogenetic dissimilarities between Nitrospira OTU0.03s across DWTPs and the associated feed GWs. The pairwise Nitrospira dissimilarities increased with distance between aquifers (r2adj = 0.88, P=0.002; Figure 5d—GW). This distance effect was significantly less when comparing Nitrospira OTUs in the PFs (r2adj = -0.11, P<2.2e−16; Figure 5d) or AFs (r2adj = -0.11, P<2.2e−16; Figure 5d) at different DWTPs.

Discussion
Microbial communities in RSFs
Proteobacteria were the dominant phylum in the PFs and remained abundant in the AFs, consistent with other, more limited, surveys of GW-fed RSFs (Gülay et al., 2014; Albers et al., 2015), and suggest their central role in RSF processes. Largest decreases in abundance from PF to AF were noted for phylotypes of the Nitrosomona- dales, an ammonium-oxidizing Beta-proteobacterial taxon, and Methylococcales, a Gamma-proteobacterial methane consuming taxon in line with the removal of their substrates from PF to AF (Macalady et al., 2013; Tatari et al., 2013; Lee et al., 2014). Even though dissolved oxygen concentrations in AFs are typically near saturation, Chloroflexi of the order Anaerolineales, assumed to be strict anaerobes (Yamada et al., 2006), were major phylotypes in AFs, suggesting anoxic microniches or another physiology. We report for the first time the presence of Gemmatimonadetes phylotypes in RSFs, but the ecology and physiology of this phylum is undescribed (DeBruyn et al., 2011).

Nitrospira was consistently the most abundant phylum in the AFs and the second most abundant phylum in the PF amplicon libraries. This high abundance was confirmed by targeted qPCR (Figure 2), although the high Nitrospira abundance caused a potential overrepresentation in the amplicon libraries (Figure 2) as observed before (Gonzalez et al., 2012). Significantly high Nitrospira occurrence, be it at lower abundance, has been seen in full-scale RSFs fed with GW (detectable Nitrospira DGGE bands (de Vet et al., 2009) and 6% of 16S rRNA amplicons (Albers et al., 2015)), fed with GW/surface water mixtures (2% of 16S rRNA amplicons (Pinto et al., 2012) and 50.4% of 16S rRNA clone libraries (White et al., 2012)) or fed with surface waters (13% to 21% of 16S rRNA amplicons (LaPara et al., 2015) and 17% of 16S rRNA clone libraries (Feng et al., 2012)). Nitrospira were also dominant in a model drinking water distribution system carrying water from a GW-fed RSF (5–27% in biofilm and 30–40% in bulk water clones; Martiny et al., 2005).

All detected Nitrospirae OTUs belonged to the Nitrospira genus (275 and 814 for PFs and AFs, respectively). The subgenus level analysis showed a broad diversity (Figure 4), suggesting a number of ecophysiology distinct Nitrospira species, as observed elsewhere (Daims et al., 2001; Pester et al., 2014). In PFs, 214 and 39 OTU0.03s belonged to Nitrospira lineage I and lineage II, respectively, whereas 39 Nitrospira OTU0.03s were of undefined lineage (Supplementary Figure S13—PF). In con- trast, in the AFs, the majority of the Nitrospira OTU0.03s (328) were of undefined lineage, with few OTU0.03s belonging to lineage II (2), III (5), IV (2) and VI (36), respectively (Supplementary Figure S13—AF).

Of all possible co-occurrence relations among the Nitrospira OTU0.03s (45.958 ± 15.228) in each DWTP
community, few were detected as significant and positive \((5590 \pm 3951, 13 \pm 8\%)\) or significant and negative \((123 \pm 95, 0.30 \pm 0.28\%)\) (Supplementary Table S6). Hence, the vast majority of relations between Nitrospira OTU\(_{0.03s}\) appear neutral instead of competitive. Taken together with the high intragenus diversity, this suggests a high degree of specialization within the Nitrospira clade, permitting the exploitation of slightly distinct niches.

The high abundance of these Nitrospira sequences in the community (Figure 2a) is puzzling given the low influent nitrite concentrations \((<0.01 \text{ mg l}^{-1}\)) and the relatively low abundance of known ammonium-oxidizing prokaryotes (Nitrospira over AOB ratios: \(45 \pm 90\) and \(210 \pm 431\) based on qPCR and amplicon libraries, respectively; Figure 2b) in the filters. Unusually high Nitrospira/AOB ratios were recently reported by others in other GW-fed filtration systems \((\text{Nitrospira: 33.4 \pm 8\%, Nitrosomonas: 3.7 \pm 2.5\%—16S rRNA amplicons; Nitzsche et al., 2015; Nitrospira: 16.9 \pm 12\%, Nitrosomonas: 3.2 \pm 2\%—16S rRNA clones; White et al., 2012), and even in surface water-fed filters (Nitrospira: 17.3\%, Nitrosomonas: 0.82\%—16S rRNA clones; Feng et al., 2012 and Nitrospira: 17 \pm 5\%, Nitrosomonas: 0.07\%—16S rRNA amplicons; LaPara et al., 2015). Differences in 16S gene copy number (Nitzsche et al., 2015), differences in specific growth rates of Nitrosomonas vs Nitrospira (Feng et al., 2012), presence of unidentified AOBs (LaPara et al., 2015), multiple energy metabolisms in Nitrospira (LaPara et al., 2015; Nitzsche et al., 2015), primer bias in PCR (Nitzsche et al., 2015) and the presence of dormant Nitrospira cells (Martiny et al., 2005) have all been invoked to explain these unexpected ratios. While all these explanation may contribute to high Nitrospira/Nitrosomonas ratios (Figure 2), the observed ratios in our study remain much higher than previously reported.

Possibly, novel physiological features support the presence of these Nitrospira phylotypes. This is in line with the co-occurrence analysis on all core taxa; significant co-occurrence patterns of Nitrospira with other taxa were exclusively negative, suggesting competitive interactions (Supplementary Figure S14). It has been documented that Nitrospira-like bacteria in activated sludge can assimilate pyruvate (Daims et al., 2001), and simple organic compounds can stimulate growth of Cand. Nitrospira defluvii and Nitrospira marina (Watson et al., 1986; Spieck et al., 2006). In addition, the genome of Cand. Nitrospira defluvii encodes pathways for the degradation and assimilation of acetate, pyruvate and formate (Lucker et al., 2010), and Nitrospira can efficiently use formate as carbon and energy source (Gruber-Dorninger et al., 2015). It is possible, then, that part of the Nitrospira abundance is supported by their growth on assimilable organic carbon substrates, present in the GW. In addition, the very recent identification of novel Nitrospira species carrying and expressing the amo operon
of all OTU$_{0.03}$s in the AF core), no close cultured isolates could be identified. Most of the unassigned OTU$_{0.03}$s in the core belong to the taxa Candida, Chloracidobacterium, Da023 (Acidobacteria), Myxococcales, Rhizobiales (a-proteobacteria), Xanthomonadales (γ-proteobacteria) and Anaerolineales (Chloroflexi). These taxa are also dominant in other GW-fed RSFs (Albers et al., 2015; Nitzsche et al., 2015). Acidobacteria are abundant in a suite of environments (Barns et al., 1999), harbor broad metabolic capabilities and can cope with nutrient limitations (Ward et al., 2009; Hartmann et al., 2015); they may be central to carbon cycling in RSFs. Myxococcales are widely distributed in terrestrial and marine environments, are often surface attached (Ganesh et al., 2014) and have a predatory lifestyle or live off polymer hydrolysis (Shimkets et al., 2006; Huntley et al., 2011); they might have a central role in recycling microbial decay products in RSFs. The physiology of Xanthomonadales is largely undescribed, but some are, similarly, known to have high extracellular hydrolytic activity (Jacobson Meyers et al., 2014).

It should be pointed out that, while an RSF core is evident at the genus level, the core taxa still retain subgenus variation across DWTPs, potentially due to dispersal limitation between the source aquifers (Gibert et al., 2009), leading to different evolutionary trajectories for any clade. Indeed, the observed distance-decay relationship for Nitrospira OTU$_{0.03}$s in the source aquifers is consistent with previous studies on aquifer communities (Franklin et al., 1999; Hug et al., 2015) and a weakened signal remains in the PF and AF Nitrospira clade.

**Ecological patterns in RSFs**

The consistent decrease in community evenness and richness from GW to filter might be due to the strong environmental changes (shift in DO level from below detection to saturation), resulting in loss of many, potentially oxygen-intolerant, taxa from GWs, and enrichment and selection of few taxa, adapted to the new environmental conditions. While the total OTU richness decreased, the OTU richness of few PF genera strikingly increased.

At the community level, OTU richness and evenness were consistently higher in the AF compared with the PF communities; this is surprising as available energy would decrease (because of microbial removal of energy sources) moving down-gradient in the DWTP (Lopato et al., 2011, 2013; Tatari et al., 2013; Lee et al., 2014) and available energy would promote richness by its positive effect on community size, per the species-energy theory (Wright 1983). Increasing community richness with decreasing energy input contradicts previous observations on natural and technical systems (Bienhold et al., 2011 and Supplementary Table S2).

Direct evidence of total community size decrease from PF to AF was difficult to obtain (small decreases were detected for DWTP2, 3 and 4 with qPCR), and suggests a significant degree of density redistribution by detachment/deposition, among other processes. The pattern of diversity increase is detected at the community, phylum, order and even at the genus level (Supplementary Table S7). OTU richness did, however, remain correlated with relative (Supplementary Table S8) abundance (Nitrospira and Nitrosomonas) within specific taxa, confirming that increasing taxon abundance increases taxon diversity (Figure 5c). These taxa were, primarily, the dominant members of the AF community.

Negative relationships between diversity and energy input have not been reported at the community level, but they have been observed at the level of specific taxa (Bienhold et al., 2011). Oligotrophy was suggested as the cause for taxa to exhibit negative relationships with energy input (Bienhold et al., 2011), and could be an explanation for the increase of abundance, and hence richness, with decreasing energy levels in RSFs. The abundance of Nitrospira as well as genera Xanthomonadales, Nkb5, Haliangium, A0839 (Proteobacteria), Acidomicrobiales, CI500 (Actinobacteria), Anaerolineales (Chloroflexi), Candida, Chloracidobacterium, Da023 (Acidobacteria) and OM190 (Planctomycetes) significantly increased with sampling location. Potentially, Nitrospira OTUs become more dominant and diverse in AFs (Figure 3b), as they are favored under low substrate conditions (Schramm et al., 1999; Attard et al., 2010; Lucker et al., 2010) (Supplementary Table S9). On the other hand, the relative abundance of core genera with putatively assigned physiology, such Nitrosomonadales, Crenothrix, Thiobacillus, Methylosoma, Hyphomicrobium and Azospira, decreases with sampling location in a DWTP, consistent with consumption of their specific energy substrates (Lopato et al., 2011, 2013; Tatari et al., 2013; Gülay et al., 2014; Lee et al., 2014).

Here, we extend our previous observation (Gülay and Smets, 2015) that RSF microbial communities display spatial heterogeneity within a single PF or AF at all DWTPs. This heterogeneity is more pronounced in AFs than PFs. In all cases, rare taxa (which are more dominant in AFs than PF) explain the dissimilarities (Table 1 and Gülay and Smets, 2015), as observed in other environments (Youssef et al., 2010). The contribution of observed community dissimilarity to differences in filter activity (Lopato et al., 2011, 2013) is uncertain, especially given the abundance of the core taxa, shared across all samples. Across DWTPs, communities are as a result less dissimilar in the RSFs than in the GWs that feed them (Supplementary Figure S12), confirming similar selective forces driving community structure (Table 1), and an identifiable resultant RSF core microbiome.
Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank K Amoako for assistance in the sample collection and C Baysal for help in script writing. We also thank G Mutlu, A Hansen and B Wagner for linguistic and other support. This research was financially supported by The Danish Council for Strategic Research (Project DW Biofilter).

Author contributions

All authors contributed to this study and the final manuscript. The original concept and sampling strategy were developed by AG, SM, HJA and BFS. Sampling was performed by AG and SM. DNA preparation and PCR experiments were performed by AG and SM. WAAS performed the pyrosequencing, supported by SJS. AG performed all the bioinformatic analyses. AG and BFS lead the analysis and interpretation of the results, the drafting, revision and completion of the manuscript. All authors read and approved the manuscript before submission.

References

Albers GN, Ellegaard-Jensen L, Harder CB, Rosendahl S, Knudsen BE, Ekelund F et al. (2015). Groundwater chemistry determines the prokaryotic community structure of waterworks sand filters. *Environ Sci Technol* 49: 839–846.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.

Attard E, Poly F, Commeaux C, Laurent F, Terada A, Smets BF et al. (2010). Shifts between *Nitrospira*- and *Nitrobacter*-like nitrite oxidizers underlie the response of soil potential nitrite oxidation to changes in tillage practices. *Environ Microbiol* 12: 315–326.

Barns SM, Takala SL, Kuske CR. (1999). Wide distribution and diversity of members of the bacterial kingdom Acidobacterium in the environment. *Appl Environ Microbiol* 65: 1731–1737.

Bell T. (2010). Experimental tests of the bacterial distance–decay relationship. *ISME J*: 4: 1357–1365.

Bienhoff C, Boetius A, Ramette A. (2011). The energy–diversity relationship of complex bacterial communities in Arctic deep-sea sediments. *ISME J*: 6: 724–732.

Burrell PC, Phalen CM, Hovanec TA. (2001). Identification of bacteria responsible for ammonia oxidation in freshwater aquaria. *Appl Environ Microbiol* 67: 5791–5800.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7: 335–336.

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

Daims H, Nielsen JL, Nielsen PH, Schleifer KH, Wagner M. (2001). *In situ* characterization of *Nitrospira*-like nitrite-oxidizing bacteria active in wastewater treatment plants. *Appl Environ Microbiol* 67: 5273–5284.

Danielopol D, Pospisil P, Rouch R. (2000). Biodiversity in groundwater: a large-scale view. *Trends Ecol Evol* 15: 223–224.

DeBruyn JM, Nixon LT, Fawaz MN, Johnson AM, Radosevich M. (2011). Global biogeography and quantitative seasonal dynamics of *Gemmataonema* in soil. *Appl Environ Microbiol* 77: 6295–6300.

de Vet WWJM, Dinkla JTI, Muyzer G, Rietveld LC, van Loosdrecht MCM. (2009). Molecular characterization of microbial populations in groundwater sources and sand filters for drinking water production. *Water Res* 43: 182–194.

Eichler S, Christen R, Hölting C, Westphal P, Bötél J, Brettar I et al. (2006). Composition and dynamics of bacterial communities of a drinking water supply system as assessed by RNA- and DNA-based 16S rRNA gene fingerprinting. *Appl Environ Microbiol* 72: 1858–1872.

Feng S, Xie S, Zhang X, Yang Z, Ding W, Liao X et al. (2012). Ammonium removal pathways and microbial community in GAC-sand dual media filter in drinking water treatment. *J Environ Sci (China)* 24: 1587–1593.

Franklin R, Taylor D, Mills A. (1999). The distribution of microbial communities in anaerobic and aerobic zones of a shallow coastal plain aquifer. *Microb Ecol* 38: 377–386.

Fredrickson JK, Onstott TC. (2001). Biogeochemical and geological significance of subsurface microbiology. In: Fredrickson JK, Fletcher M (eds), *Subsurface Microbiology and Biogeochemistry*, vol. 20. Wiley-Liss: New York, NY, USA, pp 3–37.

Fuhrman JA. (2009). Microbial community structure and its functional implications. *Nature* 459: 193–199.

Ganesh S, Parris DJ, Delong EF, Stewart FJ. (2014). Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J*: 8: 217–221.

Gaston KJ. (2000). Global patterns in biodiversity. *Nature* 405: 220–227.

Hansen M, Pjetursson B. (2011). Free, online Danish shallow geological data. *Geol Surv Den Greenl* 23: 53–56.

Gilbert J, Culver DC, Dole-Olivier M-J, Malard F, Christman MC, Dehavergre L. (2009). Assessing and conserving groundwater biodiversity: synthesis and perspectives. *Freshw Biol* 54: 930–941.

Gobet A, Quince C, Ramette A. (2010). Multivariate cutoff level analysis (MultiCoLa) of large community data sets. *Nucleic Acids Res* 38: e155.

Gonzalez JM, Portillo MC, Belda-Ferre P, Mira A. (2012). The distribution of the rare biosphere in microbial communities. *PLoS One* 7: e29973.

Griebler C, Lueders T. (2009). Microbial biodiversity in groundwater ecosystems. *Freshw Biol* 54: 649–677.

Gruber-Dorninger C, Pester M, Kitzinger K, Savio DF, Danielopol D, Pospisil P, Bötél J, Brettar I et al. (2006). Composition and dynamics of bacterial communities of a drinking water supply system as assessed by RNA- and DNA-based 16S rRNA gene fingerprinting. *Appl Environ Microbiol* 72: 1858–1872.

Gülay A, Smets BF. (2015). An improved method to set significance thresholds for β diversity testing in microbial community comparisons. *Environ Microbiol* 17: 3154–3167.
Gülay A, Tatari K, Musovic S, Mateiu RV, Albrechtsen H-J, Smets BF. (2014). Internal porosity of mineral coating supports microbial activity in rapid sand filters for groundwater treatment. *Appl Environ Microbiol* **80**: 7010–7020.

Harris JK, Kelley ST, Pace NR. (2004). New perspective on uncultured bacterial phyletogenetic division OP11. *Appl Environ Microbiol* **70**: 845–849.

Hartmann M, Frey B, Mayer J, Mäder P, Widmer F. (2015). Distinct soil microbial diversity under long-term organic and conventional farming. *ISME J* **9**: 1177–1194.

Hug LA, Thomas BC, Brown CT, Frischkorn KR, Williams KH, Tringe SG et al. (2015). Aquifer environment selects for microbial species cohorts in sediment and groundwater. *ISME J* **9**: 1846–1856.

Huntley S, Hamann N, Wegener-Feldbrügge S, Treuner-Lange A, Kube M, Reinhardt R et al. (2011). Comparative genomic analysis of fruting body formation in *Myxococcales*. *Mol Biol Evol* **28**: 1083–1097.

Jacobson Meyers ME, Sylvan JB, Edwards KJ. (2014). Extracellular enzyme activity and microbial diversity measured on seafloor exposed basalt from Loihi seamount indicate the importance of basalt to global biogeochemical cycling. *Appl Environ Microbiol* **80**: 4854–4864.

Kerr JT, Vincow ES, Currie DJ. (1998). *Lepidopteran* richness patterns in North America. *Ecocience* **5**: 448–453.

Koops H, Purkhold U, Pommerening-rö A, Timmermann G, Wagner M. (2006). The lithoautotrophic ammonia-oxidizing bacteria. In: Balows A, Trüper GG, Dworkin M, Wagner M (eds), *The Prokaryotes*. Springer: New York, New York, USA, pp 47–112.

Lopato L, Röttgers N, Binning PJ, Arvin E. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* **31**: 814–821.

LaPara TM, Hope Wilkinson K, Strait JM, Hoalski RM, Sadowsky MJ, Hamilton MJ. (2015). The bacterial communities of full-scale biologically active, granular activated carbon filters are stable and diverse and potentially contain novel ammonia-oxidizing microorganisms. *Appl Environ Microbiol* **81**: 6864–6872.

Lee CO, Boe-Hansen R, Musovic S, Smets B, Albrechtsen H-J, Binning P. (2014). Effects of dynamic operating conditions on nitrification in biological rapid sand filters for drinking water treatment. *Water Res* **64C**: 226–236.

Lopato L, Galaj Z, Delpont S, Binning PJ, Arvin E. (2011). Heterogeneity of rapid sand filters and its effect on contaminant transport and nitrification performance. *J Environ Eng* **137**: 248–257.

Lopato L, Röttgers N, Binning PJ, Arvin E. (2013). Heterogeneous nitrification in a full-scale rapid sand filter treating groundwater. *J Environ Eng* **139**: 375–384.

Lucker S, Wagner M, Maixner F, Pelletier E, Koch H, Vacherie B et al. (2010). A *Nitrospira* metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *Proc Natl Acad Sci USA* **107**: 13479–13484.

Macalady JL, Hamilton TL, Grettenberger CL, Jones DS, Tsao LE, Burgos WD. (2013). Energy, ecology and the distribution of microbial life. *Philos Trans R Soc Lond Ser B* **368**: 20120383.

Martiny AC, Albrechtsen H, Arvin E, Molin S. (2005). Identification of bacteria in biofilm and bulk water samples from a nonchlorinated model drinking water distribution system: detection of a large nitrite-oxidizing population associated with *Nitrospira* spp. *Appl Environ Microbiol* **71**: 8611–8617.

Nitschke LS, Weigold P, Lősekann-Behrens T, Kappler A, Behrens S. (2015). Microbial community composition of a household sand filter used for arsenic, iron, and manganese removal from groundwater in Vietnam. *Chemosphere* **138**: 47–59.

Pernthaler J. (2013). Freshwater microbial communities. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (eds), *The Prokaryotes*. Springer: Berlin, Heidelberg, Germany, pp 97–112.

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

Pinto AJ, Xi C, Raskin L. (2012). Bacterial community structure in the drinking water microbiome is governed by filtration processes. *Environ Sci Technol* **46**: 8851–8859.

Poindexter JS. (2006). Dimorphic prosthecate bacteria: the genera *Caulobacter, Asticcacaulis, Hyphomicrobium, Pedobacterium, Hyphomonas* and *Thiodendron*. In: Dworkin M, Falkow S, Rosenberg E, Stackebrandt E, Trüper GG, Dworkin M, Falkow S, Rosenberg E, Stackebrandt E (eds), *The Prokaryotes*. Springer: New York, New York, USA, pp 72–129.

Pommier T, Canbäck B, Riemann L, Boström KH, Simu K, Lundberg P et al. (2007). Global patterns of diversity and community structure in marine bacterioplankton. *Mol Ecol* **16**: 867–880.

Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ. (2011). Removing noise from pyrosequenced amplicons. *BMC Bioinform* **12**: 38.

Ramette A, Tiedje JM. (2007). Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem. *Proc Natl Acad Sci USA* **104**: 2761–2766.

Rivière D, Desvignes V, Pelletier E, Chaussonnerie S, Guermazi S, Weissbach J et al. (2009). Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *ISME J* **3**: 700–714.

Sanders AM, Albertsen M, Vollertsen J, Nielsen PH. (2015). The activated sludge ecosystem contains a core community of abundant organisms. *ISME J* **10**: 11–20.

Schramm A, De Beer D, Van Den Heuvel JC, Ottengraf S, Amann R. (1999). Microscale distribution of populations and activities of *Nitrosospira* and *Nitrospira* spp. along a macroscale gradient in a nitrifying bioreactor: quantification by *in situ* hybridization and the use of microsensors. *Appl Environ Microbiol* **65**: 3690–3696.

Shade A, Handelsman J. (2012). Beyond the Venn diagram: the hunt for a core microbiome. *Environ Microbiol* **14**: 4–12.

Shimkets LJ, Dworkin M, Reichenbach H. (2006). The *Myxobacteria*. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, Trüper GG, Dworkin M, Falkow S, Rosenberg E, Stackebrandt E (eds), *The Prokaryotes*. Springer: New York, New York, USA, pp 31–115.

Spieck E, Hartwig C, McCormack I, Maixner F, Wagner M, Lipski A et al. (2006). Selective enrichment and molecular characterization of a previously uncultured *Nitrospira*-like bacterium from activated sludge. *Environ Microbiol* **8**: 405–415.

Sundberg C, Al-Soud Wa, Larsson M, Alm E, Yekta SS, Svensson BH et al. (2013). 454 pyrosequencing analyses of bacterial and archaeal richness in 21 full-scale biogas digesters. *FEMS Microbiol Ecol* **85**: 612–626.
Tatari K, Musovic S, Gülay A, Dechesne A, Albrechtsen H-J, Smets BF. Density and distribution of nitrifying guilds in rapid sand filters for drinking water production (in preparation).

Tatari K, Smets BF, Albrechtsen H-J. (2013). A novel bench-scale column assay to investigate site-specific nitrification biokinetics in biological rapid sand filters. Water Res 47: 6380–6387.

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

van der Wielen PWJJ, Voost S, van der Kooij D (2009). Ammonia-oxidizing bacteria and archaea in groundwater treatment and drinking water distribution systems. Appl Environ Microbiol 75: 4687–4695.

Ward NL, Challacombe JF, Janssen PH, Henrissat B, Coutinho PM, Wu M et al. (2009). Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. Appl Environ Microbiol 75: 2046–2056.

Watson SW, Bock E, Valois FW, Waterbury JB, Schlosser U. (1986). Nitrospira marina gen. nov. sp. nov.: a chemolithotrophic nitrite-oxidizing bacterium. Arch Microbiol 144: 1–7.

White CP, Debry RW, Lytle DA. (2012). Microbial survey of a full-scale, biologically active filter for treatment of drinking water. Appl Environ Microbiol 78: 6390–6394.

Wolfe RS. (1983). Species-Energy Theory: an extension of Species-Area Theory. Oikos 41: 496–506.

Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H et al. (2006). Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the. Int J Syst Evol Microbiol 56: 1331–1340.

Youssef NH, Couger MB, Elshahed MS. (2010). Fine-scale bacterial beta diversity within a complex ecosystem (Zodletone Spring, OK, USA): the role of the rare biosphere Ahmed, N, ed. PLoS One 5: e12414.

Zhang T, Shao M-F, Ye L. (2012). Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. ISME J 6: 1137–1147.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)