Reproducible microbial community dynamics of two drinking water systems treating similar source waters.

Sarah C Potgieter[a], Ameet J Pinto[b], Minette Havenga[a], Makhosazana Sigudu[c] and Stefanus N Venter[a]

[a] *Rand Water Chair in Water Microbiology, Department of Microbiology and Plant Pathology, University of Pretoria, South Africa*

[b] *Department of Civil and Environmental Engineering, Northeastern University, Boston, USA*

[c] *Scientific Services, Rand Water, Vereeniging, South Africa*

*corresponding author: SN Venter

Email address: fanus.venter@up.ac.za

Abstract

In addition to containing higher concentrations of organics and bacterial cells, surface waters are often more vulnerable to pollution and microbial contamination with intensive industrial and agricultural activities frequently occurring in areas surrounding the water source. Therefore, surface waters typically require additional treatment, where the choice of treatment strategy is critical for water quality. Using 16S rRNA gene profiling, this study provides a unique opportunity to simultaneously investigate and compare two drinking water treatment plants and their corresponding distribution systems. The two treatment plants treat similar surface waters, from the same river system, with the same sequential treatment strategies. Here, the impact of treatment and distribution on the microbial community within and between each system was compared over an eight-month sampling campaign. Overall, reproducible spatial and temporal dynamics within both DWTPs and their corresponding DWDSs were observed. Although source waters showed some dissimilarity in microbial community structure and composition, pre-disinfection treatments (i.e. coagulation, flocculation, sedimentation and filtration) resulted in highly similar microbial communities between the filter effluent samples. This indicated that the same treatments resulted in the development of similar microbial communities. Conversely, post-disinfection (i.e. chlorination and chloramination) resulted in increased dissimilarity between disinfected samples from the two systems, showing alternative responses of the microbial community to disinfection. Lastly, it was observed that within the distribution system the same dominant taxa were selected where samples increased in similarity with increased residence time. Although, differences were found between the two systems, overall treatment and distribution had a similar impact on the microbial community in each system. This

study therefore provides valuable information on the impact of treatment and distribution on the drinking water microbiome.

Keywords: drinking water treatment; drinking water distribution; disinfection; microbial community dynamics; Illumina MiSeq.

Highlights

- Source waters show some dissimilarity in microbial community.
- Treatment processes increases similarity and selects for the same dominant taxa.
- Differential response to chlorination causing increased dissimilarity and variation.
- Stabilisation of DWDS microbial community through selection of same dominant taxa.
- Microbial community dynamics are reproducible between the two systems.

Abbreviations

DWDS, drinking water distribution system; DWTP, drinking water treatment plant; DADA2, Divisive Amplicon Denoising Algorithm; ASV, amplicon sequence variant; AMOVA, analysis of molecular variance; MRA, mean relative abundance; PCoA, Principal coordinate analysis; ANOVA, One-way analysis of variance.

1. Introduction

Drinking water is a vital resource and is therefore one of the most closely monitored and strictly regulated resources. Rapid urbanisation, agricultural expansion, and climate change have resulted in the alteration of natural water systems (specifically surface water) that now challenge the performance of water treatment facilities (Delpla *et al*., 2009; Poitelon *et al*., 2010). Treatment operations are designed to reduce microbial concentrations and limit microbial growth in drinking water distribution systems (DWDS). Nevertheless, drinking water treatment plants (DWTPs) are typically biodiverse and harbour complex microbial ecosystems (Bruno *et al*., 2018). Following coagulation, flocculation, sedimentation, and filtration, modern DWTPs employ multi-barrier treatment processes that demonstrate microbial removal/disinfection efficacies (i.e., chlorination and/or chloramination, ozonation and UV-disinfection) to ensure the production of high-quality drinking water. The choice of treatment strategy is a fundamental decision, which is highly site specific and based on the characteristics of the source water (Prest *et al*., 2016).

Previous studies have shown that the drinking water microbiome is considerably impacted by the choice of treatment strategy and distribution (Pinto *et al*., 2012; Bautista-de los Santos *et al*., 2016; Prest *et al*., 2016; Potgieter *et al*., 2018; Zhang *et al*., 2017). Here, treatment and distribution processes may be considered as ecological disturbances implemented sequentially on the microbiome continuum within drinking water (Zhang *et al*., 2017). In many European countries, disinfection is not used as the final step in treatment. In such cases, water treatment may involve multiple barriers and extensive biofiltration with the focus on nutrient removal

(Hammes *et al*., 2010; Lautenschlager *et al*., 2013). However, in the cases where disinfection is used, it is well established that it significantly reduces microbial numbers and alters the microbial community composition and abundance (Gomez-Alvarez *et al*., 2012; Hwang *et al*., 2012; Prest *et al*., 2016; Potgieter *et al*., 2018). Despite these methods to reduce/limit microbial numbers, microbes persist and form indigenous inhabitants of the distribution system despite low nutrient levels and disinfectant residuals. Here, finished drinking water typically maintains cell concentrations between 10^3 and 10^5 cells/mL (Hammes *et al*., 2010; Lautenschlager *et al*., 2013; Liu *et al.,* 2013; Gillespie *et al*., 2014; Nescerecka *et al*., 2014).

The persistence and growth of microorganisms in DWDSs are responsible for many of the problems associated with the drinking water distribution systems. Microbial growth is often responsible for nitrification in chloraminated systems (Kirmeyer *et al*., 1995; Wilczak *et al.,* 1996; Wang *et al*., 2014b), increased disinfectant demand (Vasconcelos *et al*., 1997) and through biofilm formation, they promote the deterioration of pipe surfaces through microbial mediated corrosion (MIC) and can also harbour potential pathogens (Boe-Hansen *et al.,* 2002; Berry *et al*., 2006; Ling *et al*., 2018). Furthermore, microbial water quality can continue to deteriorate during distribution as a result of bacterial growth due to insufficient disinfectant residuals (Fish and Boxall, 2018), changes in water supply and consumption or stagnation (Ling *et al*., 2018), seasonal fluctuations (Pinto *et al.,* 2014; Potgieter *et al*., 2018) and the influence of mixing of different water sources (Pinto *et al*., 2014; Nescerecka *et al*., 2018).

Factors influencing the drinking water microbiome are undeniably site specific due to unique DWTP and DWDS configurations, water sources, water quality and operational practices (Pinto *et al*., 2012). Studies have demonstrated the site-specific impacts of treatment and distribution on the drinking water microbiome and compared these across multiple DWTP's and DWDS (Roeselers *et al.,* 2015; Gulay *et al*., 2016). While useful at drawing generalised trends on the impact of specific treatment processes and/or distribution system configurations, these cross system comparisons are often linked to differing source waters. Source water type typically has a significant impact on the microbial community composition and structure and thereby potentially masks the true impact of treatment and distribution. This presents a gap in the literature, as to our knowledge, no study has investigated similar treatment and distribution of two source waters in the same large-scale DWDS.

This study presents unique insights into the systematic comparison between two DWTPs (treating similar source waters, originating from the same river system) and their corresponding distribution systems. More specifically, the two similar source waters are subjected to the same sequential treatment strategies within the two different DWTPs and resulting treated water is distributed in within the same large-scale DWDS, although across diverging lines. We hypothesize that the same treatment strategies and similar distribution of the drinking water will result in the development of similar microbial communities. Using 16S rRNA gene profiling, the current study aims to investigate the reproducibility of the microbial community dynamics in these two drinking water systems. The scope of this study involved an eight-month sampling campaign where samples were collected monthly from

corresponding sample locations from the two systems. The specific objectives were to; (i) investigate the difference in community composition and structure of the two source waters, (ii) determine the effect of treatment and distribution in shaping the microbial communities in the two systems, (iii) identify the dominant taxa responsible for differences in community assemblages and (iv) evaluate the differential distribution patterns between the two systems.

2. Materials and methods

2.1 Site description

This research presents a unique opportunity to systematically compare two drinking water systems (System R and S), which include two treatment plants (treating similar source waters) and their corresponding distribution networks, all forming part of the same large-scale DWDS and under the operation of the same drinking water utility (Fig. 1A). As a whole, this drinking water utility covers a vast network, stretching over 3056 km of pipeline and covering 18,000 km^2. It supplies on average 4800 million liters per day to approximately 12 million people within large metropolitan and local municipalities as well as mines and industries. The source water is drawn primarily from a river and dam system via two drinking water treatment plants (R_DWTP and S_DWTP), which abstract, purify and pump 98% (approximately 4320 ML/d) of the total water supplied by the utility. The R_DWTP (river intake pumping site) treats source water from the river downstream of the dam and the S_DWTP treats source water from a canal directly from the dam.

Treatment of the source waters in both DWTPs consists of the same conventional purification steps (Fig. 1B). Briefly, source water in both DWTPs is dosed with

polyelectrolyte coagulants with low lime for coagulation and flocculation, with no need for pH correction after sedimentation. Although in some months in System R, a combination of polyelectrolyte and silica lime was used in coagulation and flocculation (Table S1). In those cases, following sedimentation, the pH of the alkaline water is adjusted to near neutral by bubbling CO_2 gas followed by filtration through rapid gravity sand filters. Finally, the filter effluent is dosed with chlorine gas is bubbled into carriage water to be dosed into the main water for disinfection.

The total chlorine at sites following chlorination varies between 1.0 mg/L and 2.5 mg/L after 20 min contact time. Chlorinated water leaving both DWTPs is again dosed with chloramine (approximately 2 mg/L) at a secondary disinfection boosting stations. For the purpose of this study chlorinated water originating from the R_DWTP was followed to a booster station, which produces approximately 1 100 ML/d of chloraminated water, serving predominately the northwest area of the distribution system (R_DWDS). Chlorinated water originating from the S_DWTP was also followed to another booster station, producing approximately 700 ML/d of chloraminated water to the eastern parts of the distribution system (S_DWDS) (Fig. 1A). Here within the chloraminated sections of the DWDS, monochloramine residuals vary on average between 0.8 mg/L in the autumn and 1.5 mg/L in the spring. These monochloramine residual concentrations don’t differ significantly between the two systems and range from approximately 2 mg/L immediately following chloramination to 1.4 mg/L at the end points in the DWDSs. Further details on range of physical-chemical parameters for both systems were obtained from the utility (Table S1, S2A and S2B).

2.2 Sample collection and processing

Samples were collected for 8 months (February 2016 – September 2016) from corresponding study sites from two DWTPs (R_DWTP and S_DWTP) and their associated DWDS networks (R_DWDS and S_DWDS) (Fig. 1A). Study sites within the two DWTPs included source water (SW), filter inflow (FI, i.e. water entering the rapid sand filter following coagulation, flocculation, sedimentation and carbonation), filter bed media (FB), filter effluent (FE, i.e. following filtration) and chlorinated water leaving the treatment plant (CHLA). Within the two DWDS sections study sites included chlorinated water entering the secondary disinfection booster station before chloramination (CHLB), chloraminated water leaving the booster station (CHM) and chloraminated bulk water at two points with the DWDSs (DS1 and DS2, respectively) (Fig. 1A and 1B). Within the two DWTPs, 1 L of source water, 4 L of filter inflow, 8 L of filter effluent and 8 L of bulk water were collected. Typically, for samples collected directly after disinfection, 8 – 16 L of bulk water was collected. Collected water samples were filtered to harvest microbial cells followed by phenol:chloroform DNA extraction as described by Potgieter *et al*., 2018.

To obtain microbial biomass from the filter bed media samples, 10 g of filter media was mixed with 50 ml extraction buffer (i.e., 0.4 g/L EGTA, 1.2 g/L TRIS, 1 g/L peptone and 0.4 g/L N-dodecyl-N, N dimethyl-3-amminio-1-propanesulfonate) followed by sonication for 1 min to remove the microbial biomass attached to sand particles (Camper *et al*., 1985). After sonication, the aqueous phase was filtered through a SterivexTM-GP 0.22 µm polycarbonate membrane filter unit (Merck Milipore, South Africa) followed by phenol:chloroform DNA extraction, as with the water samples.

2.3 Sequencing and data processing

Extracted genomic DNA from samples were sent to the Department of Microbiology and Immunology, University of Michigan Medical School (Ann Arbor, USA) for the sequencing of the V4 hypervariable region of 16S rRNA gene using the Illumina MiSeq platform. Sequencing was performed using a paired-end sequencing approach described by Kozich *et al.* (2013), resulting in 250 nucleotide long paired reads. All raw sequence data have been deposited with links to BioProject accession number PRJNA529765 in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/).

A total of 181 samples were successfully sequenced. Sequence analysis of these samples was performed using the Divisive Amplicon Denoising Algorithm, DADA2 (Callahan *et al*., 2016). Full amplicon workflow included sequence filtering, dereplication, inferring sample composition, chimera identification and removal, merging of paired-end reads and construction on a sequence table. Initial trimming and filtering of reads followed standard filtering parameters described for Illumina MiSeq 2x250 V4 region of the 16S rRNA gene (https://benjjneb.github.io/dada2/tutorial.html) where reads with ambiguous bases were removed (maxN=0), the maximum number of “expected errors” was defined (maxEE=2) and reads were truncated at the first instance of a quality score less than or equal to truncQ (truncQ=2). Dereplication was performed where identical sequences are combined into “unique sequences” while maintaining the corresponding abundance of the number of reads for that unique sequence. The core sample inference algorithm was applied to dereplicated data and forward and reverse reads were merged together to obtain fully denoised sequences (Callahan

et al., 2016). Merged reads were then used to construct an amplicon sequence variant (ASV) table (Callahan *et al*., 2017), chimeras were identified and removed and taxonomic assignments were called using the SILVA reference database (https://www.arb-silva.de) through the DADA2 taxonomy assignment script.

2.4 Microbial community analysis

Resulting ASV table was imported into the mothur software package (v 1.35.1) (Schloss *et al*., 2009) and the shared sequences between sample locations from the two DWTPs and corresponding DWDSs as well as the unique sequences within each sample location were calculated using the venn function in mothur. Furthermore, alpha diversity measures (i.e., richness, Shannon Diversity Index and Pielou's evenness) were calculated using the summary.single function in mothur with the parameters, subsampling=1263 (sample with the least ab=mount of sequences) and iters=1000 (1000 subsampling of the entire dataset). Due to subsampling, 10 samples were excluded from the analyses and Good's coverage estimates were calculated to assess whether sufficient number of sequences were retained for each sample after subsampling. This indicated that subsampling at a library size of 1263 retained the majority of the richness for all samples (i.e., average Good's coverage = 95.84 ± 0.02%). One-way analysis of variance (ANOVA) (Chambers *et al.,* 1992) and post-hoc Tukey Honest Significant Differences (HSD) test were performed in R (http://www.R-project.org) using the stats package (R Core Team, 2015) to determine the statistical significance between spatial and temporal groupings within the alpha diversity.

Temporal and spatial variabilities in the microbial community structure and membership were calculated using beta diversity assignment methods, i.e. Jaccard and Bray-Curtis distances as well as phylogenetic placement method, i.e. weighted and unweighted UniFrac. Bray-Curtis and weighted UniFrac (as calculated based on presence/absence and abundance data) were used for the analysis of community structure as pair-wise dissimilarity between selected samples, whereas Jaccard and unweighted UniFrac (calculated based on presence and absence data) were used to infer community membership. Bray-Curtis and Jaccard distances were calculated using the dist.shared function in mothur with the parameters, subsampling=1263 and iters=1000. Weighted and unweighted UniFrac distances were calculated through the construction of a phylogenetic tree with representative sequences using the clearcut command in mothur also with the parameters subsampling=1263 and iters=1000 (Evans *et al*., 2006; Lozupone *et al*., 2011).

Pairwise Analysis of Molecular Variance (AMOVA) was performed using the amova function in mother on all beta diversity matrices, to determine the effect of sample groupings based on DWDS sample location, DWDS section and season (Excoffier, 1993; Anderson, 2001). Beta diversity metrics and metadata files containing sample location, sample type, disinfection type and season were imported into R (http://www.R-project.org) for statistical analysis. Principal-coordinate analyses (PCoA) using Bray-Curtis and Jaccard distances was performed using the phyloseq package (McMurdie and Holmes, 2013). All plots were constructed using the ggplot2 package (Wickham, 2009).

3. Results

3.1 Microbial community composition of the two systems

Overall, 10,012 ASV's were identified, constituting 4,921,399 sequences. Taxonomic classification of these ASV's revealed that bacteria dominated the microbial community (mean relative abundance, MRA 98.74 ± 0.02% across all samples) followed by archaea (MRA 1.04 ± 0.01%). Overall, comparisons between corresponding samples from Systems R and S showed similar microbial community compositions. Although the two source waters harboured the same bacterial phyla, these phyla differed in relative abundance. Water originating from the river (R_SW) had higher mean relative abundance of *Proteobacteria* (MRA: 41.04 ± 6.98%) than the source water originating from the dam (S_SW) (MRA: 23.99 ± 9.36%). Conversely, S_SW showed higher relative abundances of *Actinobacteria* (MRA: 31.14 ± 2.03 %) than R_SW (MRA: 20.69 ± 7.14%). *Bacteroidetes* showed moderately high relative abundance and remained constant between the two source waters (i.e. R_SW MRA: 13.49 ± 8.78% and S_SW MRA: 12.29 ± 3.62%) (Fig. 2A and Table S3).

Between the two varying source waters, only 22.51% of the ASV's identified were shared (i.e., 711 ASV's) (Fig. S1). These shared ASV's made up 6.93% of the total sequences and 47.67% and 28.75% of the total ASV abundance in R_SW and S_SW, respectively. Of these shared ASV's, approximately 30% had a MRA of ≥ 0.05% across the respective source water samples. However, these top 30% of abundant ASV's were found to differ in relative abundance depending on the source water origin. Overall, ASV_2 (*Actinobacteria,* family *Sporichthyaceae*) was found to be dominant in both R_SW and S_SW with MRA of 4.82 ± 2.78% and 8.42 ± 0.97%,

respectively. The relative abundance of other dominant ASV's differed between the two source waters. Here, ASV_57 and ASV_71 (both *Proteobacteria,* genus *Pseudomonas*), ASV_7 (*Actinobacteria,* family *Sporichthyaceae*), and ASV_15 (Archaea, *Thaumarchaeota*) showed increased MRA across R_SW samples with MRA of 3.10 ± 1.51%, 3.30 ± 3.90%, 2.95 ± 0.83% and 2.76 ± 1.60%, respectively. Within S_SW samples, ASV_5 (*Actinobacteria,* family *Sporichthyaceae*), ASV_15 (Archaea, *Thaumarchaeota*), ASV_7 (*Actinobacteria,* family *Sporichthyaceae*) and ASV_91 (*Proteobacteria*, genus *Hydrogenophaga*) showed increased relative abundance with MRA of 7.63 ± 1.40%, 6.40 ± 2.55%, 4.28 ± 0.87% and 2.47 ± 1.61%, respectively. Throughout both DWTPs (i.e., including SW, FI, FB and FE samples), *Proteobacteria*, *Actinobacteria* and *Bacteroidetes* were dominant (MRA: 30.51 ± 10.63% and 26.31 ± 8.53% and 16.62 ± 6.83%, respectively). The microbial community composition of both DWTPs was highly diverse and included other dominant phyla (i.e. MRA greater than 1%) i.e., *Acidobacteria* (MRA: 9.05 ± 5.04%), *Cyanobacteria* (MRA: 2.89 ± 2.29%), *Verrucomicrobia* (MRA: 2.62 ± 1.32%) and *Planctomycetes* (MRA: 2.07 ± 2.41%) (Fig. 2A and Table S3).

However, a change in community composition was observed following chlorination in both systems. Here, on average, a decrease in *Actinobacteria* and *Bacteroidetes* was observed (MRA: 6.25 ± 10.19% and 2.26 ± 4.44%, respectively). This corresponded to increases in *Planctomycetes* and *Cyanobacteria* (MRA: 12.78 ± 15.32% and 4.07 ± 4.05, respectively). However, the community composition differed between chlorinated samples from the two systems (R_CHLA and S_CHLA). Samples from R_CHLA showed increased relative abundance of *Planctomycetes* (MRA: 17.45 ± 18.38%) and decreased relative abundance of

Proteobacteria (MRA: 25.50 ± 7.33%), whereas the converse was observed in S_CHLA samples (*Planctomycetes* MRA: 6.24 ± 6.81% and *Proteobacteria* MRA: 36.51 ± 15.86%). Also, samples from R_CHLB showed higher abundances of *Actinobacteria* and *Bacteroidetes* (MRA: 8.73 ± 10.30% and 12.30 ± 18.96%, respectively) than S_CHLB samples (*Actinobacteria* MRA: 1.41 ± 0.06% and *Bacteroidetes* MRA: 0.86 ± 0.59%). Similar to S_CHLA samples, S_CHLB samples also showed an increase in *Proteobacteria* (MRA: 40.03 ± 18.84%) as well as decreased *Planctomycetes* (MRA: 2.77 ± 4.20%) (Fig. 2A).

Following chloramination, the community composition became more similar again between corresponding samples from the two systems. In contrast to DWTP samples, within the chloraminated section of the DWDSs, CHM samples showed an increase in the relative abundance of *Planctomycetes* (R_CHM MRA: 21.63 ± 25.11% and S_CHM MRA: 21.98 ± 15.06%) and *Proteobacteria* specifically in R_CHLB (MRA: 47.86 ± 27.98%). Both R_CHM and S_CHM also showed a decrease in *Actinobacteria* (MRA: 3.07 ± 2.15% and 2.55 ± 2.71%, respectively). *Proteobacteria* reached its highest relative abundances in the distribution system samples DS1 and DS2 (MRA: 65.92 ± 13.98% and 70.09 ± 2.57%, respectively) samples in both systems (Fig. 2A).

Due to the dominance of *Proteobacteria* in CHM, DS1 and DS2 samples from both systems, investigations into the relative abundance of proteobacterial classes revealed that *Alpha-* and *Gammaproteobacteria* were the most dominant (MRA of 28.71 ± 14.65% and 29.25 ± 10.32%, respectively) (Fig. 2B). However, within *Gammaproteobacteria,* the order *Betaproteobacteriales* showed high mean relative

abundance of 16.37 ± 9.24% across CHM, DS1 and DS2 samples. Interestingly, the dominance of *Alphaproteobacteria* and *Betaproteobacteriales* varied between DS1 and DS2 samples from R_DWDS versus the corresponding samples from S_DWDS. Within R_DS1, *Betaproteobacteriales* dominated with a MRA of 21.85 ± 21.70% followed by *Alphaproteobacteria* with a MRA of 13.88 ± 8.78%. Conversely, S_DS1 samples were dominated by *Alphaproteobacteria* with a MRA of 49.16 ± 21.19% followed by *Betaproteobacteriales* with a MRA of 18.42 ± 12.13%. The same dominance of *Alphaproteobacteria* was observed in S_DS2 samples (i.e. *Alphaproteobacteria* MRA: 43.01 ± 18.66% and *Betaproteobacteriales* MRA: 23.74 ± 15.38%). Lastly, R_DS2 samples showed shared dominance between *Alphaproteobacteria* and *Betaproteobacteriales* with similar MRA (i.e. *Alphaproteobacteria* MRA: 29.54 ± 18.10% and *Betaproteobacteriales* MRA: 25.60 ± 15.81%) (Fig. 2B).

3.2 Spatial trends in abundance of dominant bacterial taxa across both systems

An investigation into the spatial trends of the most abundant ASV's revealed that only 14 (i.e., 0.14% of total ASV's and constituting 34.75% of all sequences) had a mean relative abundance of ≥1% across both systems (Fig. S2, Table S4A and S4B). Gülay *et al.* (2016) describes dominant or core taxa as those shared taxa with a relative abundance >1% were. Therefore, these 14 ASV's could be considered as the dominant core taxa across both systems. Throughout both DWTPs the same ASV's dominated, showing similar distributions across SW, FI, FB, and FE samples (Fig. 3). This included other ASV's that were abundant in both DWTPs (mean relative abundance of ≥1%) but decreased in relative abundance in the DWDSs.

These other dominant ASV's included members of *Thaumarchaeaota* (ASV_15), *Actinobacteria* (ASV_17 and ASV_21), *Bacteroidetes* (ASV_20) and *Betaproteobacteriales* (ASV_27). Specifically, the ASV's that dominated both DWTPs were ASV_2, ASV_5 and ASV_7 (all belonging to *Actinobacteria*, family *Sporichthyaceae*), ASV_4 (*Acidobacteria*, family *Holophagaceae*), ASV_8 (*Proteobacteria*, family *Burkholderiaceae*) and ASV_13 (*Proteobacteria*, family *Pheatobacter*). Here, the mean relative abundance of these dominant ASV's was higher in the S_DWTP than in R_DWTP (Tables B2A and B2B). Some ASV's, i.e. ASV_1 (*Proteobacteria*, genus *Methylobacterium*), ASV_6 (*Proteobacteria*, genus *Nitrosomonas*), ASV_12 (*Acidobacteria*) and ASV_41 (*Proteobacteria,* class *Alphaproteobacteria*), while abundant in the DWDSs, where not detected across both R_DWTP and S_DWTP samples.

Treatment may select for the same dominant ASV's in both systems, although the dominance of these ASV's differed in abundance between R_FE and S_FE. Filter effluent samples (R_FE and S_FE) shared 36.04% of the total ASV between the two groups, constituting 56.89% and 49.57% of the ASV's in R_SW and S_SW, respectively. These shared ASV's included the dominant ASV's in both R_FE and S_FE and could be traced back to both source waters. Dominant ASV's shared between the two source waters, i.e., ASV_2, ASV_5 and ASV_7 (all belonging to *Actinobacteria,* family *Sporichthyaceae*) with MRA across SW samples of 6.62 ± 2.54%, 6.50 ± 1.60% and 3.62 ± 0.94%, respectively, remained dominant in the FE samples (MRA ASV_2: 8.95 ± 0.84%, ASV_5: 5.20 ± 1.01% and ASV_7: 5.23 ± 0.63%).

Following chlorination, the MRA and distribution of these dominant ASV's changed significantly (Fig. 3). At CHLA sites, all ASV's showing high MRA in the DWTPs decreased significantly and ASV_3 (*Planctomycetes*, family *Gemmataceae*) and ASV_41 (*Alphaproteobacteria*) increased in both R_CHLA and S_CHLA, although the MRA of these two ASV's was higher in R_CHLA than in S_CHLA. However, S_CHLA also showed small increases in the MRA of ASV_1 (*Proteobacteria*, genus *Methylobacterium*), ASV_6 (*Proteobacteria*, genus *Nitrosomonas*) and ASV_12 (*Acidobacteria*). A difference between CHLB samples from the two systems was also observed. Here, in S_CHLB all ASV's decreased except for ASV_41 (*Alphaproteobacteria*), whereas in R_CHLB the MRA of most ASV's was higher.

Chloramination and distribution generally resulted in the increase in the MRA of ASV's that were absent or had low MRA in the DWTPs (Fig. 3). These ASV's included ASV_1 (*Proteobacteria*, genus *Methylobacterium*), ASV_3 (*Planctomycetes*, family *Gemmataceae*), ASV_6 (*Proteobacteria*, genus *Nitrosomonas*) and ASV_12 (*Acidobacteria*) and ASV_30 (*Planctomycetes*, genus *Planctomyces*). Amplicon sequence variants ASV_10 (*Proteobacteria*, genus *Pseudomonas*) and ASV_11 (*Proteobacteria*, genus *Sphingomonas*) maintained a generally consistent MRA across both systems. Interestingly, S_DS1 and S_DS2 distribution sample sites showed very similar abundances and distribution of dominant taxa, whereas in R_DS1 and R_DS2 this pattern was not observed as ASV's differed in abundance. Throughout both systems these dominant ASV's showed the same distribution across all samples, although their abundances differed between the two systems.

3.3 Reproducible spatial trends in alpha diversity of two parallel drinking water systems

Both source waters were significantly richer (average observed taxa: 256 ± 44) and more diverse (Shannon Diversity Index: 4.40 ± 0.48 and Inverse Simpson Diversity Index: 43.08 ± 16.27) when compared to all other samples within their corresponding DWTPs and DWDSs (Fig. 4 and Table S5). Source water originating from the dam (S_SW) was more rich (average observed taxa: 304 ± 17) than the source water originating from the river (R_SW) (208 ± 70) and the differences in richness between these source water samples were found to be significant ($p < 0.05$) based on one-way analysis of variance, ANOVA and post-hoc Tukey Honest Significant Differences (HSD) test. However, the diversity and evenness between the two source waters were found to be similar [(S_SW, average Shannon Diversity index 4.66 ± 0.10, average Inverse Simpson Diversity Index 40.69 ± 5.40 and average Pielou's evenness 0.81 ± 0.01) R_SW, average Shannon diversity index 4.30 ± 0.55, average Inverse Simpson Index 39.58 ± 16.05 and average Pielou's evenness 0.82 ± 0.03)] and the differences in diversity and evenness between source water samples were not significant (ANOVA; $p > 0.001$).

Significant differences in alpha diversity measures were predominately observed between spatial groupings (i.e., between different sample locations) (ANOVA; richness: $F_{ST} = 19.67$, $p < 0.05$, Shannon Diversity Index: $F_{ST} = 9.78$, $p < 0.05$, Inverse Simpson Diversity Index: $F_{ST} = 15.64$, $p < 0.05$ and Pielou's evenness: $F_{ST} = 4.79$, $p < 0.05$). Overall, DWTP samples, from both systems, were more rich and diverse than those in the DWDSs. Although, richness and diversity consistently decreased along treatment processes (excluding filter bed samples (FB)) reflecting

the changes in the community caused by each treatment step (Fig. 4 and Table S5). The same trends in all alpha diversity measures were observed for all corresponding sample comparisons between both System R and S. Specifically, decreases in richness and diversity were observed in the FI samples following coagulation, flocculation and sedimentation (average observed taxa: 191 ± 40; Shannon Diversity Index: 4.16 ± 0.31 and Inverse Simpson Diversity Index: 31.82 ± 9.93), in FE samples following sand filtration (average observed taxa: 160 ± 26; Shannon Diversity Index: 3.93 ± 0.30 and Inverse Simpson Diversity Index: 25.42 ± 8.28) and finally, the most significant decrease in CHLA samples following chlorination (average observed taxa: 68 ± 45; Shannon Diversity Index: 2.88 ± 0.59 and Inverse Simpson Diversity Index: 11.88 ± 7.52).

However, following chloramination (i.e., sites CHM, DS1 and DS2), all alpha diversity measures increased as the distance from the site of chloramination increased (CHM; average observed taxa: 83 ± 32; Shannon Diversity Index: 3.06 ± 0.64 and Inverse Simpson Diversity Index: 12.79 ± 7.68, DS1; average observed taxa: 102 ± 44; Shannon Diversity Index: 3.06 ± 0.89 and Inverse Simpson Diversity Index: 14.25 ± 9.06 and DS2; average observed taxa: 123 ± 42; Shannon Diversity Index: 3.32 ± 0.66 and Inverse Simpson Diversity Index: 14.52 ± 8.12). In terms of evenness, samples within the two DWTPs were the most even (Pielou’s evenness: 0.80 ± 0.03) and following chlorination and chloramination evenness decreased (Pielou’s evenness: 0.70 ± 0.10), albeit not significantly (Fig. 4 and Table S5).

3.4 Reproducible spatial trends in microbial community structure and membership in both systems

Beta diversity metrics indicated that the two source waters were dissimilar in both community membership (i.e., Jaccard: 0.84 ± 0.06 and unweighted UniFrac: 0.71 ± 0.06) and community structure (i.e., Bray-Curtis: 0.71 ± 0.10 and weighted UniFrac: 0.47 ± 0.13). These dissimilarity values were also found to be statistically significant (AMOVA, $F_{ST} \leq 3.04$, $p < 0.001$, depending on the beta diversity measure). Further, R_SW samples showed increased temporal variability in community structure compared to S_SW samples. Here, consecutive temporal R_SW samples within the 8 month study period showed increased dissimilarity in community structure (i.e., Bray-Curtis: 0.65 ± 0.10 and weighted UniFrac: 0.44 ± 0.12) compared to S_SW samples (i.e., Bray-Curtis 0.52 ± 0.08 and weighted UniFrac: 0.27 ± 0.06) (Fig. 5).

Pairwise beta diversity comparisons between consecutive samples from each system showed similar spatial trends (Fig. 6). Here, treatment and distribution have the same impact on the microbial community structure and membership in both systems. The filter bed samples from both DWTPs were shown to significantly different from both the filter inflow (AMOVA: $F_{ST} \leq 5.07$, $p < 0.001$) and filter effluent (AMOVA: $F_{ST} \leq 6.51$, $p < 0.001$). Although, in sample comparisons from both DWTPs, the microbial community became increasingly more similar from source water through treatment and filtration where the microbial community in filter bed and filter effluent are approximately 40 – 60% similar in community structure (Bray-Curtis: 0.62 ± 0.08, weighted UniFrac: 0.42 ± 0.10) and 30 – 40% similar in community membership (Jaccard: 0.72 ± 0.04 and unweighted UniFrac: 0.62 ± 0.04).

Comparisons involving disinfection (chlorination and chloramination) showed increased dissimilarity in both community structure and membership. The microbial community become significantly more dissimilar following chlorination where the microbial communities between filter effluent (FE) and bulk water immediately after chlorination (CHLA) were approximately 80 – 85% dissimilar in community structure (Bray-Curtis: 0.88 ± 0.18, weighted UniFrac: 0.81 ± 0.18) and membership (Jaccard: 0.89 ± 0.16 and unweighted UniFrac: 0.78 ± 0.15) (AMOVA: $F_{ST} \leq 18.22$, $p < 0.001$ depending on the beta diversity measure). Conversely, chlorinated locations CHLA and CHLB increased in similarity in community structure (Bray-Curtis: 0.62 ± 0.19, weighted UniFrac: 0.49 ± 0.18) and were found not to be significantly different (AMOVA: $p \leq 0.697$). Again, following chloramination, the microbial communities between chlorinated (CHLB) and chloraminated water (CHM) increased significantly in dissimilarity in community structure (Bray-Curtis: 0.84 ± 0.13, weighted UniFrac: 0.71 ± 0.13) and membership (Jaccard: 0.89 ± 0.05 and unweighted UniFrac: 0.77 ± 0.05) (AMOVA: $F_{ST} \leq 4.09$, $p < 0.001$ depending on the beta diversity measure) in both systems. Lastly, following chloramination, microbial communities within the two DWDS showed converse spatial trends. In System R, the microbial community structure in DS1 samples showed increased similarity with CHM samples (Bray-Curtis: 0.72 ± 0.17, weighted UniFrac: 0.57 ± 0.16) and increased in dissimilarity with DS2 samples (Bray-Curtis: 0.78 ± 0.09, weighted UniFrac: 0.55 ± 0.09) (AMOVA: $F_{ST} \leq 2.55$, $p < 0.001$ depending on the beta diversity measure). Conversely, in System S, a marginal increase in dissimilarity was observed in community structure between CHM and DS1 samples and a significant increase in similarity in microbial community structure between DS1 and DS2 samples (Bray-

Curtis: 0.54 ± 0.21, weighted UniFrac: 0.42 ± 0.17) (AMOVA: $p \leq 0.655$). However, in both systems CHLB, CHM, DS1 and DS2 samples remain constant and unchanged in microbial community membership.

Principle coordinate analysis (PCoA) of all samples from both systems revealed clustering of all DWTP samples regardless of which DWTP system they originated (Fig. 7A). This correlated with observed similarity in pairwise beta diversity comparisons between DWTP from both systems. However, no clear clustering was observed for all DWDS samples, which also correlated with observed increases temporal and spatial variability in DWDS samples from both systems. Individual PCoAs of both DWTPs (Fig. 7B) and DWDSs (Fig. 7C) based on Bray-Curtis distances revealed limited clustering of samples based on the system they originated from. However, the PCoA ordination of DWTPs samples showed a shift in between samples as they moved through the DWTP (Fig. 7B). Samples from different locations showed some clustering but also showed overlap with consecutive samples sites. With the exception of source water and filter inflow. Complete overlap between filter inflow and filter effluent samples were observed. Although clustering was not pronounced, a shift or succession in samples was also observed in DWDS samples where chlorinated samples and those samples immediately following chloramination grouped closer together (Fig. 7C). Samples from both distribution systems showed little or no concise clustering, which may be due to temporal variations within each location.

3.5 Temporal trends were similar across both drinking water systems

The same temporal trends were observed in both community membership (Jaccard and unweighted UniFrac) and structure (Bray-Curtis and weighted UniFrac) in samples within DWTP (FI, FB and FE) from both System R and S. Within these sample sites, increased dissimilarity between samples 6 months apart was observed, indicating seasonal variations, although the eight month sample period was insufficient to observe complete seasonal trends. However, the changes in temporal dissimilarity were marginal, indicating general temporal stability within the microbial communities of DWTP samples and samples towards the end of the DWDS for both systems. DWTP samples were observed to be more temporally stable as pair-wise comparisons between consecutive months within each sample location were less dissimilar in community membership (i.e., Jaccard: 0.62 ± 0.08 and unweighted UniFrac: 0.53 ± 0.06) and structure (i.e., Bray-Curtis: 0.48 ± 0.11 and weighted UniFrac: 0.32 ± 0.12).

Interestingly, samples following disinfection (i.e., CHLA, CHLB and CHM), from both systems, indicated increased temporal variability within each sample location with increased dissimilarity in community membership (i.e., Membership Jaccard: 0.87 ± 0.05 and unweighted UniFrac: 0.75 ± 0.06) and structure (Bray-Curtis: 0.72 ± 0.16 and weighted UniFrac: 0.59 ± 0.18) (Fig. 5). Specifically, chlorinated samples R_CHLA and S_CHLA showed converse temporal trends, where R_CHLA samples 6 months apart increased in similarity in both community structure and membership. CHLB samples from both systems then showed similar temporal trends but also increased in similarity as the months between samples increased. Similarly, this trend was also observed in the microbial community structure and membership of

S_CHM samples, although the changes in dissimilarity within these samples were marginal. Samples within the DWDS (DS1 and DS2) showed consistent temporal trends where samples from both systems increased in dissimilarity 6 months apart. However, temporal variability remained high within DS1 and DS2 samples although lower than samples following disinfection where pair-wise comparisons between consecutive months within each sample location were dissimilar in community membership (i.e., Jaccard: 0.80 ± 0.06 and unweighted UniFrac: 0.67 ± 0.06) and structure (i.e., Bray-Curtis: 0.69 ± 0.13 and weighted UniFrac: 0.52 ± 0.13) (Fig. 5).

Temporal trends were also observed when focusing on individual ASV's, however no single ASV was present at every time point across all samples. Therefore, temporal trends were observed for ASV's present at all time points within all DWTP (SW, FI, FB and FE) samples, chlorinated samples (CHLA and CHLB) and DWDS samples (CHM, DS1 and DS2) separately. Furthermore, ASV's were considered dominant if they obtained a MRA ≥ 1% across specific sample groups. Here, dominant ASV's that occurred at all time points in both DWTPs were identified as ASV_2 (*Actinobacteria,* family *Sporichthyaceae*), ASV_7 (*Actinobacteria,* family *Sporichthyaceae*), ASV_15 (*Thaumarchaeota,* genus *Ca.* Nitrosoarchaeum), ASV_24 (*Cyanobacteria,* genus *Cyanobium*) and ASV_27 (*Betaproteobacteriales*, genus *Ca.* Methylopumilus). The temporal variation of these ASV's across the two DWTPs was generally similar, however variability was observed that was specific to each individual ASV and specific sample location. This variability in temporal trends for each ASV and sample site was also observed when considering ASV's [ASV_3 (*Planctomycetes*, family *Gemmataceae*) and ASV_54 (*Planctomycetes*)] of moderate abundance (MRA 1% < and > 0.1% across DWTP samples).

No clear temporal trends were observed in sample sites following chlorination (CHLA and CHLB) from both systems, as no single ASV was present at all time points and temporal trends were observed to be highly variable across dominant and moderately abundant ASV's. Although not present at all time points within CHLA and CHLB locations, ASV_3 (*Planctomycetes*, family *Gemmataceae*), ASV_40 and ASV_41 (both *Alphaproteobacteria)* were identified as the dominant ASV's. Here, their temporal variation across the eight months occurred in a converse relationship between the two systems indicating high temporal variability at these locations.

Interestingly, the same temporal trends were observed in dominant and moderately abundant ASV's in DWDS samples. Here, only ASV_3 (*Planctomycetes*, family *Gemmataceae*) was observed to be dominant and in all CHM, DS1 and DS2 samples from both systems. This ASV showed the same temporal trends in CHM, DS1 and DS2 sample sites in both systems. Furthermore, ASV_1 (*Proteobacteria*, genus *Methylobacterium*) and ASV_6 (*Proteobacteria*, genus *Nitrosomonas*) were found to be dominant and in all DS1 and DS2 samples. These two ASV's showed highly similar trends in both systems, indicating increased temporal stability towards the end of both DWDSs. The same general temporal trends were also observed between the two systems in moderately abundant ASV's present at all time points in the DWDS samples, i.e., ASV_36 (unclassified) and ASV_83 (*Nitrospira*), although more variable.

3.6 Disinfection increased microbial community dissimilarity across the two drinking water systems.

Beta diversity comparisons between corresponding samples from System R and S were calculated in line with the layout of treatment and distribution as well as for corresponding months (Fig. 8 and Table S6). Following coagulation, flocculation, sedimentation and pH adjustment, the filter inflow samples between the two DWTPs (R_FI and S_FI) became significantly more similar in community structure (Bray-Curtis: 0.49 ± 0.11 and weighted UniFrac: 0.31 ± 0.09) and membership (Jaccard: 0.66 ± 0.10 and unweighted UniFrac: 0.56 ± 0.10). Similarly, following sand filtration, R_FE and S_FE sample comparisons maintained the same level of similarity (Bray-Curtis: 0.48 ± 0.13, weighted UniFrac: 0.34 ± 0.16, Jaccard: 0.58 ± 0.08 and unweighted UniFrac: 0.50 ± 0.04) and were found not to be significantly different (AMOVA for both FI and FE sample comparisons: $F_{ST} \leq 1.72$, $p \leq 0.077$ depending on beta diversity measure). Decreased beta diversity dissimilarity values indicated greater stability in both microbial community structure and membership in DWTP samples, specifically FI and FE samples. Although, filter bed microbial communities (R_FB and S_FB) were showed increased dissimilarity in community structure (Bray-Curtis: 0.66 ± 0.08 and weighted UniFrac: 0.42 ± 0.13) and membership (Jaccard: 0.75 ± 0.05 and unweighted UniFrac: 0.63 ± 0.03) (AMOVA: $F_{ST} \leq 2.57$, $p < 0.001$).

Conversely, samples immediately after chlorination (R_CHLA and S_CHLA) showed an increase in dissimilarity in both community structure (Bray-Curtis: 0.72 ± 0.20 and weighted UniFrac: 0.69 ± 0.24) and membership (Jaccard: 0.92 ± 0.04 and unweighted UniFrac: 0.82 ± 0.07) (Fig. 8). Similar dissimilarity in community

structure (Bray-Curtis: 0.72 ± 0.24 and weighted UniFrac: 0.51 ± 0.32) and membership (Jaccard: 0.93 ± 0.06 and unweighted UniFrac: 0.82 ± 0.10) was observed between R_CHLB and S_CHLB samples. Although these samples increased in dissimilarity between the two systems, the values were not statistically significant (AMOVA: $F_{ST} \leq 1.18$, $p \leq 0.492$ for both CHLA and CHLB comparisons depending on the beta diversity measure). This may be due to the high level of temporal variability observed between individual comparisons between these sample locations. The same may be true for chloraminated samples, as samples R_CHM and S_CHM showed similar dissimilarity values as CHLA and CHLB as well as high temporal variability between individual samples comparisons (Bray-Curtis: 0.74 ± 0.21, weighted UniFrac: 0.67 ± 0.16, Jaccard: 0.82 ± 0.06 and unweighted UniFrac: 0.69 ± 0.05 and AMOVA: $F_{ST} \leq 1.19$, $p \leq 0.274$). Lastly, chloraminated sites with the DWDS (R_DS1 and S_DS1) maintained increased dissimilarity (Bray-Curtis: 0.72 ± 0.23, weighted UniFrac: 0.67 ± 0.11, Jaccard: 0.83 ± 0.08 and unweighted UniFrac: 0.72 ± 0.04). Following the bulk water further down the DWDS (R_DS2 and S_DS2) the samples increased slightly in similarity in community structure (Bray-Curtis: 0.67 ± 0.17 and weighted UniFrac: 0.56 ± 0.10) but remained the same in community membership (Jaccard: 0.83 ± 0.08 and unweighted UniFrac: 0.72 ± 0.04). This dissimilarity between DS1 and DS2 samples from the two systems was observed to be significantly different (AMOVA: $F_{ST} \leq 5.09$, $p < 0.001$) (Fig. 8).

4. Discussion

4.1 Dissimilarity in microbial community observed between similar source waters

Consistent with previous studies, the microbial composition of source waters and DWTPs was dominated by *Proteobacteria* and *Actinobacteria, Bacteroidetes, Acidobacteria, Cyanobacteria, Planctomycetes* and *Verrucomicrobia.* These phyla are known to be common in freshwater (i.e. rivers, lakes and dams) (Newton *et al*., 2011; Martinez-Garcia *et al*., 2012) and DWTPs (Poitelon *et al*., 2010; Kwon *et al*., 2011; Pinto *et al*., 2012; Zeng *et al*., 2013; Lautenschlager *et al* 2014; Lin *et al*., 2014) and are capable of utilising a variety of substrates. The source water microbial communities were highly diverse and significantly richer than the microbial communities in all other DWTP and DWDS samples. This observation was unsurprising, as source water comprised of surface water that was obtained from a large temperate, nutrient-rich eutrophic system and not subjected to prior physical or chemical treatment. It is important to note that the two source water sites are part of the same river system and are not independent from each other. However, the two source waters showed high dissimilarity in microbial community structure and membership; this dissimilarity may arise from geographical and hydrological differences. The microbial community of the source water originating from the river may be subjected to strong hydrological conditions such as runoff and increased flow rates during heavy rainfall events before the source water is channelled into the DWTP (Prathumratana *et al*., 2008; Delpla *et al*., 2009). As a result, this source water also showed higher temporal variability where samples over the eight month period increased in dissimilarity. Conversely, the microbial community of the source water originating from the dam may experience stagnation and was more temporally

stable. This difference in hydrological parameters and temporal variability between the two source waters may then translate to the occurrence of different rare or low abundant taxa specific to each source water, resulting in the increased dissimilarity in microbial community structure and membership (Shade *et al*., 2014).

4.2 Treatment shapes the core microbial community

Corresponding samples from the two DWTPs also showed similar abundances of the dominant phyla, indicating stability of dominant groups across the two treatment plants. Although the abundance of these phyla differed across sample sites, their dominance was maintained throughout all DWTP samples, suggesting that the impact of coagulation, flocculation and sedimentations on the microbial community at the phylum level was relatively small (Lin *et al*., 2014). Microbial community richness and diversity consistently decreased with consecutive treatment operations in both systems. This suggests that a decrease in microbial relative abundance, which typically occurs during the treatment processes (Hammes *et* al., 2008; Prest *et al.,* 2014; Lin *et al*., 2014; Wang *et al*., 2014b), also resulted in changes in the diversity of the microbial community. Drinking water treatment typically consists of sequential treatment operations that operate continuously to deliver microbially safe drinking water and although connected, each independent treatment step introduces potential physicochemical variability, thereby impacting the microbial community. The microbial community between the two DWTP showed increased similarity in samples following coagulation, flocculation, sedimentation and carbonation as well as after sand filtration (Kwon *et al*., 2009; Poitelon *et al*., 2010; Lin *et al*., 2014). The pair-wise comparisons between filter inflow and filter effluent samples from both systems also revealed increased similarity between samples. Furthermore, the

microbial community from DWTP samples from both systems showed the same temporal trends and were consistently more temporally stable than the microbial communities in other samples. These findings suggest that these treatment operations have the similar impact on the microbial community membership and structure from the two DWTPs and the similarities in their design and operational parameters leads to shared dominant DWTP microbial communities (Gulay *et al*., 2016).

The filter bed microbial community showed increased dissimilarity compared to the communities within filter inflow and filter effluent samples. Rapid gravity sand filters receive continuous inputs from the source water and this input may vary depending on the temporal and spatial dynamics of the source (Gulay *et al*., 2010). Therefore, the establishment and integration of bacteria into the biofilm community of the sand filter is significantly influenced by the physicochemical properties and microbial community of the source waters. Gulay *et al*. (2010) suggested that heterogeneity between the microbial communities of sand filters from different DWTP could be explained by rare taxa and the development of differing biofilm communities on the filter bed. This may be the case in this study as the two filter beds shared only 29.22 % of the total richness. Furthermore, backwashing of the filter beds with finished chlorinated water may also contribute to the dissimilarity between the two filter beds, which corresponds to the dissimilarity observed between chlorinated samples from the two systems (Liu *et al*., 2012; Liao *et al*., 2015b). Although, across the two DWTPs, microbial communities were more similar between filter bed samples than the source waters that feed them, confirming the selective forces of treatment driving community structure and the presence of dominant taxa. The filter beds may

have differed from each other but the filter effluent from both DWTP were increasingly similar. The influence of sand filtration was limited, presenting similar phylum/class level microbial community composition in the filter inflow and filter effluent bulk water samples. It is likely that bacteria from the bulk water attach to sand filters and establish and integrate themselves in the biofilm community of the sand filter as in these systems sand filter beds are backwashed with finished chlorinated water (Lin *et al*., 2014). This, together with an increase in the number of shared ASV's between the two filter effluents (36.04%), indicates that conditions in the filter beds were sufficiently similar to have the same effect on the resulting effluent and the selection of dominant taxa in both systems.

Core taxa dominant in DWTP locations suggests that treatment drives selection of the community assemblage. Core taxa within the DWTPs comprised primarily of *Actinobacteria* (*Sporichthyaceae*)*, Acidobacteria* (*Halophagaceae*), *Alphaproteobacteria, Gammaproteobacteria* and *Betaproteobacteriales* (*Rhizobiales, Phreatobacter*). These groups have previously found to be ubiquitous in DWTPs (Pinto *et al*., 2012; Lautenschlager *et al*., 2013; Zeng *et al*., 2013; Liao *et al*., 2015a). In this study, these taxa were observed to be dominant in the source waters and showed continued dominance throughout DWTP samples. These results indicated that the source water may seed the drinking water system and plays a role in shaping the microbial community within the treatment plant. Three of the top dominant ASV's in DWTP samples from both systems were identified as *Actinobacteria,* family *Sporichthyaceae*. This is consistent with other DWTP studies where this group of bacteria have adapted to the selective pressures of treatment and are competitive under low nutrient conditions (Zeng *et al*., 2013; Lin *et al*.,

2014). *Actinobacteria* have been observed to be a highly abundant phyla in freshwater lakes due to their free living style and are able to use a wide range of easily degradable organic carbon compounds (Gulay *et al*., 2010; Newton *et al*., 2011). *Acidobacteria* were also observed to be dominant in DWTP samples and are known to harbour a broad range of metabolic capabilities as well as cope with limited nutrient availability (Ward *et al*., 2009). Members of the order *Rhizobiales* are also ubiquitous in freshwater systems and are commonly found in DWTPs, where they are presumed to use a wide range of substrates (Pinto *et al*., 2012; Lautenschlager *et al*., 2013; Zeng *et al*., 2013; Lin *et al*., 2014).

4.3 Communities are impacted differently by chlorination

Chlorination significantly reduces bacterial cell concentrations and has a substantial influence on community composition and structure (Eichler *et al*., 2006; Poitelon *et al*., 2010; Wang *et al*., 2014a; Lin *et al*., 2014; Prest *et al*., 2016; Potgieter *et al*., 2018). While pre-chlorination microbial communities were similar between the two DWTPs, the microbial community composition and structure in both systems were highly dissimilar post-chlorination. This dissimilarity was also observed on a temporal scale, where chlorinated samples showed differing temporal trends and increased temporal variability. Here, the system level dynamics at the point of disinfection may be stronger than the temporal dynamics and therefore drives the microbial community composition and structure at these locations (Potgieter *et al*., 2018).

Significant differences were observed in microbial composition between corresponding chlorinated samples (CHLA and CHLB) between systems, indicating

high system level variability at these locations. This instability in microbial community composition in chlorinated samples has been previously documented where proteobacterial population shifts occurred due to changes in chlorine residual concentrations (Mathieu *et al*., 2009). The microbial composition within DWDS samples was consistent with that of previous studies (Bautista-de los Santos *et al*., 2016; Potgieter *et al*., 2018). It is important to acknowledge here that through disinfection, cell numbers are significantly impacted and a considerable fraction of bacteria are inactivated. However, without absolute abundance measurements and viability assays in this study, the proportion of dead cells or extracelluar DNA is unknown. Therefore, while the observed changes in the dominance of phyla and overall community composition do not address absolute abundance or viability (Sakcham *et al*., 2019), considering that the same treatment strategies are applied in both systems, we estimate cell concentrations would not differ significantly and therefore comparisons of the microbial community composition and structure could be made between corresponding samples.

Interestingly, the microbial communities following chlorination from both systems were significantly different from each other in community membership and structure, suggesting that the microbial community's response to the disturbance/stress of disinfection was different. Chlorine is non-specific in its action of reducing bacterial cell concentrations and therefore communities may be altered differently in response to chlorination. Although the ecological role of low abundant and/or rare taxa is not well understood, these taxa may act as a potential microbial seedbank when conditions change (e.g. after chlorination). Following chlorination, different taxa specific to each location may persist as they may exhibit differential resistance

to disinfection (Poitelon *et al*., 2010; Shade *et al*., 2014; Chiao *et al*., 2014). In addition, a change in substrate concentrations following disinfection may provide rare taxa alternative niches for remaining bacteria once disinfected residuals have been depleted (Shade *et al*., 2014; El-Chakhtoura *et al*., 2015; Prest *et al*., 2016). Within disinfected samples, *Planctomycetaceae* showed a significant increase in abundance, potentially suggesting greater resistance to chlorine and chloramine exposure and rapid recovery. The persistence of certain dominant ASV's in disinfected samples suggests that these taxa may exhibit a variety of functional traits that allow their survival in a range of environments from the eutrophic surface water at the source to the nutrient limited conditions and disinfection stress of the disinfected water in the DWDS (Pinto *et al*., 2012).

4.4 Potential steady state obtained through distribution

Following chloramination and through distribution, *Proteobacteria* and *Planctomycetes* dominated. However, the dominance of the proteobacterial classes *Alphaprotebacteria* and *Gammaproteobacteria* (order *Betaproteobacteriales*) across CHM, DS1 and DS2 samples differed between the two systems. The high abundance of *Proteobacteria* in drinking water systems is well documented (Lautenschlager *et al*., 2013; Liu *et al*., 2013; Bautista-de los Santos *et al*., 2016) and the inconsistency in the relative abundance of *Alpha-* and *Betaproteobacteria* (now reclassified as the order *Betaproteobacteriales* within *Gammaproteobacteria*) across different drinking water microbiomes as well as between different stages within a single system has been observed (Mathieu *et al*., 2009; Prest *et al*., 2014; Proctor and Hammes, 2015). This difference in dominance of proteobacterial classes in DWDS samples correlated with the abundance of the proteobacterial

classes in the chlorinated samples from each system. This may be attributed differences in disinfectant residual concentrations between the two sections of the DWDS (Hwang *et al*., 2012), despite the fact that there was no significant difference in monochloramine residual concentrations at the end of both systems. The difference in dominance of the two proteobacterial classes may also be result of site specific dynamics within each DWDS section, such as pipe material, pipe age and biofilm formation (Wang *et al*., 2014a; Prest *et al*., 2016).

Water distribution conditions can have a considerable impact on the drinking water microbiome (Prest *et al*., 2016). Various factors influence the microbial dynamics within the DWDS including pipe material, hydraulic conditions, residence time, water temperature and disinfectant residual concentrations. In this study, two DWDS lines originating from the two DWTP showed some dissimilarity i.e., approximately 60 – 70% dissimilarity in community structure and 70 – 80% dissimilar in community membership, where disinfectant residual concentration and water temperatures did not differ between the two lines.

Therefore, the observed dissimilarity between the two distribution lines may be accredited to the differential response of the microbial community to chlorination. However, an increase in similarity was observed in locations towards the end of the DWDS (DS2 samples). Through distribution, water is continually seeded by similar microbial communities over time thereby selecting for the same dominant taxa through similarities in pipe material, residence times, hydraulic conditions and operation practices contributing to site specific taxa and biofilms. This increase in similarity in community membership and structure with increasing residence time in

the DWDS was more pronounced in samples from summer and autumn. Here, elevated water temperatures in summer months may affect the bacterial community composition and structure by positively influencing the growth kinetics and competition processes of specific bacterial species in each section of the DWDS (Prest *et al*., 2016).

In DWDS samples, the high abundance of a *Methylobacterium*-like ASV is consistent with other studies as *Methylobacterium* has been found to be ubiquitous in cloraminated DWDS as planktonic cells or forming part of biofilms (Gallego *et al*., 2005; Gomez-Alvarez *et al*., 2012 and 2016; Wang *et al*., 2013). Furthermore, as observed by Potgieter *et al*. (2018), samples from summer and autumn months (specifically February) showed increased abundance of a *Nitrosomonas*-like ASV, which became dominant in DS2 samples, specifically in S_DS2 samples. Here, the addition of chloramine as a secondary disinfectant has been shown to support the growth of nitrifying bacteria in DWDS. The long residence time and associated lower disinfectant residual concentrations, together with the release of ammonia through disinfection decay results in increased numbers of nitrifiers and therefore potential nitrification (Wang *et al*., 2014b).

5. Conclusions

The drinking water microbiome can be considered as a continuum that travels from the source water through treatment and distribution systems, where different disturbances (through treatment and disinfection) are intentionally introduced to produce microbially safe drinking water. This study allowed for a unique opportunity to compare the effect of the same treatment strategies (disturbances) on similar

source waters as well as the distribution of treated water on the drinking water microbiome in a large-scale system. Here, we were able to show the reproducible spatial and temporal dynamics of two DWTPs and their corresponding DWDS sections within the same drinking water system. Treatment (i.e., pre-disinfection) of the two source waters produced highly similar microbial communities in the filter effluent, suggesting that similarities in design and operational parameters of the two DWTPs results in the development of similar microbial communities. However, the dissimilarity observed in the microbial community between post-disinfection samples from the two systems highlighted the differential impact of disinfection, where the response to disinfection differed between the two systems. Lastly, the influence of distribution was also observed, where certain dominant taxa were selected. Dissimilarities in microbial community throughout distribution may arise from initial differences in the source waters and the differential response to chlorination, leading the presence of site specific rare/low abundant taxa. In summary, although there are dissimilarities inherent to each location, treatment and distribution had the same impact on the microbial community in each system and may select for the same dominant species. Therefore, using 16S rRNA gene community profiling, this study provides valuable information regarding the influence of treatment and distribution on the drinking water microbiome.

Acknowledgements

This research was funded and supported by Rand Water, Gauteng, South Africa through the Rand water Chair in Water Microbiology at the University of Pretoria. Sarah Potgieter would also like to acknowledge the National Research Foundation (NRF) for additional funding. Furthermore, the authors would like to acknowledge

the Centre for Microbial Systems Molecular Biology Lab, University of Michigan, USA for their services in Illumina MiSeq sequencing.

APPENDIX A. SUPPLEMENTARY DATA

References

1. Anderson, M. J., (2001). A new method for non-parametric multivariate analysis of variance. *Australian Ecology*. 26 (1): 32e46.
2. Bautista-de los Santos, Q. M., Schroeder, M. C., Sevillano-Rivera, M. C., Sungthong, R., Ijaz, U. Z., Sloan, W. T. and Pinto, A. J. (2016). Emerging investigators series: microbial communities in full-scale drinking water distribution systems – a meta-analysis. *Environmental Science: Water Research and Technology.* doi: 10.1039/c6ew00030d.
3. Berry, D., Xi, C. and Raskin, L. (2006). Microbial ecology of drinking water distribution systems. *Current Opinion Biotechnology.* 17: 297-302.
4. Boe-Hansen, R., Albrechtsen, H. J., Arvin, E. and JØrgensen, C. (2002). Bulk water phase and biofilm growth in drinking water at low nutrient concentrations. *Water Research.* 36: 4477-4486.
5. Bruno, A., Sandionigi, A., Bernasconi, M., Panio, A., Labra, M. and Casiraghi, M. (2018). Changes in the drinking water microbiome: effects of water treatments along the flow of two drinking water treatment plants in an urbanized area, Milan (Italy). *Frontiers in Microbiology.* 9: 2557.
6. Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A. and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods*. 13(7): 581.

7. Callahan, B. J., McMurdie, P. J. and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal*. 11(12): 2639.
8. Camper, A. K., Lechevallier, M. W., Broadaway, S. C. and McFeters, G. A. (1985). Evaluation of procedures to desorb bacteria from granular activated carbon. *Journal of Microbiological Method.* 3: 187-198.
9. Chambers, J. M., Freeny, A. and Heiberger, R. M. (1992). *Analysis of variance; designed experiments.* Chapter 5 of *Statistical Models in S* eds J. M. Chambers and T. J. Hastie, Wadsworth and Brooks/Cole.
10. Chaio, T., Clancy, T. M., Pinto, A., Xi, C. and Raskin, L. (2014). Differential resistance of drinking water bacterial populations to monochloramine disinfection. *Environmental Science and Technology.* 48: 4038-3-37.
11. Delpla, I., Jung, A.V., Baures, E., Clement, M. and Thomas, O. (2009). Impacts of climate change on surface water quality in relation to drinking water production. *Environment International*. 35(8): 1225-1233.
12. Eichler, S., Christen, R., Höltje, C., Westphal, P., Bötel, J., Brettar, I., Mehling, A. and Höfle, M. G. (2006). Composition and dynamics of bacterial communities of a drinking water supply system as assessed by RNA-and DNA-based 16S rRNA gene fingerprinting. *Applied and Environmental Microbiology*. 72(3): 1858-1872.
13. El-Chakhtoura, J., Prest, E., Saikaly, P., van Loosdercht., Hammes, F. and Vrouwenvedler, H. (2015). Dynamics of bacterial communities before and after distribution in a full-scale drinking water network. *Water Research.* 74: 180-190.
14. Evans, J., Sheneman, L. and Foster, J.A. (2006). Relaxed neighbour-joining: a fast distance- based phylogenetic tree construction method. *Journal of Molecular Evoltion*. 62: 785e792.

15. Excoffier, L., (1993). Analysis of Molecular Variance (AMOVA) Version 1.55. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
16. Fish, K. and Boxall, J. (2018). Biofilm Microbiome (Re) Growth Dynamics in Drinking Water Distribution Systems Are Impacted by Chlorine Concentration. *Frontiers in Microbiology*. 9: 2519.
17. Gallego, V., Garcı ́a, M.T. and Ventosa, A. (2005). Methylobacterium hispanicum sp. nov. and methylobacterium aquaticum sp. nov., isolated from drinking water. *International Journal of Systematic Evolution in Microbiology*. 55: 281e287.
18. Gillespie, S., Lipphaus, P., Green, J., Parsons, S., Weir, P., Juskowiak, K., Jefferson, B., Jarvis, P. and Nocker, A. (2014). Assessing microbiological water quality in drinking water distribution systems with disinfectant residual using flow cytometry. *Water Research.* 65: 224-234.
19. Gomez-Alvarez, V., Revetta, R. P. and Santo Domingo, J. W. (2012). Metagenomic analysis of drinking water receiving different disinfection treatments. *Appied and Enivronmental Microbiology.* 78(17): 6095-6102.
20. Gomez-Alvarez, V., Pfaller, S., Pressman, J.G., Wahman, D.G. and Revetta, R.P., (2016). Resilience of microbial communities in a simulated drinking water distribution system subjected to disturbances: role of conditionally rare taxa and potential implications for antibiotic-resistant bacteria. *Environmental Science: Water Research and Technology.* 2(4): 645-657.
21. Gülay, A., Musovic, S., Albrechtsen, H. J., Al-Soud, W. A., Sørensen, S. J. and Smets, B. F. (2016). Ecological patterns, diversity and core taxa of microbial communities in groundwater-fed rapid gravity filters. *The ISME journal*. 10(9): 2209.

22. Hammes, F., Berney, M., Wang, Y., Vital, M., Koster, O. and Egli, T. (2008). Flow-cytometric total bacterial cell counts as a descriptive microbiological parameter for drinking water treatment processes. *Water Research*. 44(17): 4868-4877.

23. Hammes, F., Berger, C., Koster, O. and Egli, T. (2010). Assessing biological stability of drinking water without disinfectant residuals in a full-scale water supply system. *Journal of Water Supply: Research and Technology-AQUA.* 59: 31-40.

24. Hwang, C., Ling, F., Andersen, G. L., LeChevallier, M. W. and Liu, W. (2012). Microbial community dynamics of an urban drinking water distribution system subjected to phases of chloramination and chlorination treatments. *Applied and Environmental Microbiology.* 78(22): 7856-7865.

25. Kirmeyer, G. J., Odell, L. H., Jacagelo, J., Wilczak, A. and Wolfe, R. L. (1995). Nitrification occurrence and control in chloraminated water systems. Denver CO: AWWA Research Foundation and America Water Works Association. .

26. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., Schloss, P. D. (2013). Development of a dual-index strategy and curation pipeline for analyzing amplicon- sequencing data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology.* 79: 5112e5120.

27. Kwon, S., Moon, E., Kim, T.S., Hong, S. and Park, H.D. (2011). Pyrosequencing demonstrated complex microbial communities in a membrane filtration system for a drinking water treatment plant. *Microbes and Environments*. 26(2): 149-155.

28. Lautenschlager, K., Hwang, C., Ling, F., Liu, W. T., Boon, N., Köster, O., Vrouwenvelder, H., Egli, T. and Hammes, F. (2013). A microbiology-based multi-parametric approach towards assessing biological stability in drinking water distribution networks. *Water Research*. 47: 3015-3025.

29. Lautenschlager, K., Hwang, C., Ling, F., Lui, W. T., Boon, N., Köster, O., Egli, T. and Hammes, F. (2014). Abundance and composition of indigenous bacterial communities in a multi-step biofiltration-based drinking water treatment. *Water Research.* 62: 40-52.
30. Liao, X., Chen, C., Wang, Z., Chang, C.H., Zhang, X. and Xie, S. (2015a). Bacterial community change through drinking water treatment processes. *International Journal of Environmental Science and Technology*. 12(6): 1867-1874.
31. Liao, X., Chen, C., Zhang, J., Dai, Y., Zhang, X. and Xie, S. (2015b). Operational performance, biomass and microbial community structure: impacts of backwashing on drinking water biofilter. *Environmental Science and Pollution Research*. 22(1): 546-554.
32. Lin, W., Yu, Z., Zhang, H. and Thompson, I. P. (2014). Diversity and dynamics of microbial communities at each step of treatment plant for potable water generation. *Water Research*. 52: 218-230.
33. Ling, F., Whitaker, R., LeChevallier, M. W. and Liu, W. T. (2018). Drinking water microbiome assembly induced by water stagnation. *The ISME journal*. 12(6), p.1520.
34. Liu, B., Gu, L., Yu, X., Yu, G., Zhang, H. and Xu, J. (2012). Dissolved organic nitrogen (DON) profile during backwashing cycle of drinking water biofiltration. *Science of the Total Environment*. 414: 508-514.
35. Liu, G., Verbeck, J. Q. J. C. and Van Dijk, J. C. (2013). Bacteriology of drinking water distribution systems: an integral and multidimensional review. *Applied and Environmental Microbiology*. 97: 9265-9276.

36. Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., Knight, R. (2011). UniFrac: an effective distance metric for microbial community comparison. *The ISME Journal*. 5(2):169e172.

37. Martinez-Garcia, M., Swan, B. K., Poulton, N. J., Gomez, M. L., Masland, D., Sieracki, M. E. and Stepanauskas, R. (2012). High-throughput single-cell sequencing identifies photoheterotrophs and chemoautotrophs in freshwater bacterioplankton. *The ISME journal*. 6(1): 113.

38. Martínez-Hidalgo, P. and Hirsch, A. M. (2017). The nodule microbiome: N2-fixing rhizobia do not live alone. *Phytobiomes.* 1(2): 70-82.

39. Mathieu, L., Bouteleux, C., Fass, S., Angel, E. and Block, J. C. (2009). Reversible shift in the α-, β- and γ-proteobacteria populations of drinking water biofilms during discontinuous chlorination. *Water Research*. 43(14): 3375-3386.

40. McMurdie, P.J. and Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One.* 8(4): e61217.

41. Nescerecka, A., Rubulis, J., Vital, M., Juhna, T and Hammes, F. (2014). Biological instability in a chlorinated drinking water distribution network. *PloS One*. 9: e96354.

42. Nescerecka, A., Juhna, T. and Hammes, F. (2018). Identifying the underlying causes of biological instability in a full-scale drinking water supply system. *Water Research.* 135: 11-21.

43. Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D. and Bertilsson, S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiology and Molecular biology reviews*. 75(1): 14-49.

44. Pinto, A. J., Xi, C. and Raskin, L. (2012). Bacterial community structure in the drinking water microbiome is governed by filtration processes. *Environmental Science and Technology.* 46: 8851-8859.

45. Pinto, A., Schroeder, J., Lunn, M., Sloan, W. and Raskin, L. (2014). Spatial-temporal survey and occupancy-abundance modelling to predict bacterial community dynamics in the drinking water microbiome. *mBIO.* 5(3): e01135-14.

46. Poitelon, J. B., Joyeux, M., Welté, B., Duguet, J. P., Prestel, E. and DuBow, M. S. (2010). Variations of bacterial 16S rDNA phylotypes prior to and after chlorination for drinking water production from two surface water treatment plants. *Journal of Industrial Microbiology and Biotechnology*. 37(2): 117-128.

47. Potgieter, S., Pinto, A., Sigudu, M., Du Preez, H., Ncube, E. and Venter, S. (2018). Long-term spatial and temporal microbial community dynamics in a large-scale drinking water distribution system with multiple disinfectant regimes. *Water Research.* 139: 406-419.

48. Prathumratana, L., Sthiannopkao, S. and Kim, K.W. (2008). The relationship of climatic and hydrological parameters to surface water quality in the lower Mekong River. *Environment International*. 34(6): 860-866.

49. Prest, E. I., El-Chakhtoura, J., Hammes, F., Saikaly, P.E., Van Loosdrecht, M.C.M. and Vrouwenvelder, J. S. (2014). Combining flow cytometry and 16S rRNA gene pyrosequencing: a promising approach for drinking water monitoring and characterization. *Water Research*. 63: 179-189.

50. Prest, E. I., Hammes, F., van Loosdrecht, M. C. M. and Vrouwenvelder, J. S. (2016). Biological stability of drinking water: controlling factors, methods and challenges. *Frontiers in Microbiology.* 7(45): doi: 10.3389/fmicb.2016.00045.

51. Proctor, C. R. and Hammes, F. (2015). Drinking water microbiology – from measurement to management. *Current Opinion in Biotechnology*. 33: 87-94.
52. R Core Team (2015). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
53. Roeselers, G., Coolen, J., van der Wielen, P. W. J. J., Jaspers, M. C., Atsma, A., deGraaf, B. and Schuren, F. (2015). Microbial biogeography of drinking water: patterns in phylogenetic diversity across space and time. *Environmental Microbiology*. 17(7): 2505-2514.
54. Sakcham, B., Kumar, A. and Cao, B., 2019. Extracellular DNA in monochloraminated drinking water and its influence on DNA-based profiling of a microbial community. *Environmental Science and Technology Letters*.
55. Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*. 75(23): 7537e7541.
56. Shade, A., Jones, S. E., Caporaso, J. G., Handelsman, J., Knight, R., Fierer, N., Gilbert, J. A. (2014). Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *mBio ASM*. 5(4): 1e9.
57. Vasconcelos, J. L., Rossman, L. A., Grayman, W. M., Boulos, P. F. and Clark, R. M. (1997). Kinetics of chlorine decay. *Journal – American Water Works Association.* 89(7): 54-65.
58. Wang, H., Pryor, M. A., Edwards, M. A., Falkinham, J. O. and Pruden, A. (2013). Effect of GAC pre-treatment and disinfectant on microbial community structure and opportunistic pathogen occurrence. *Water Research*. 47: 5760–5772. doi:

10.1016/j.watres.2013.06.052

59. Wang, H., Masters, S., Edwards, M. A., Falkinham, J. O., and Pruden, A. (2014a). Effect of disinfectant, water age, and pipe materials on bacterial and eukaryotic community structure in drinking water biofilm. *Environmental Science and Technology*. 48: 1426–1435. doi: 10.1021/es402636u

60. Wang, H., Proctor, C. R., Edwards, M. A., Pryor, M., Santo Domingo, J. W., Ryu, H., et al. (2014b). Microbial community response to chlorine conversion in a chloraminated drinking water distribution system. *Environmental Science and Technology*. 48: 10624–10633. doi: 10.1021/es502646d.

61. Ward, N. L., Challacombe, J. F., Janssen, P. H., Henrissat, B., Coutinho, P. M., Wu, M., Xie, G., Haft, D. H., Sait, M., Badger, J. and Barabote, R. D. (2009). Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. *Applied and Environmental Microbiology*. 75(7): 2046-2056.

62. Wickham, H. (2009). ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York. http://ggplot2.org.

63. Wilczak, A., Jacangelo, J. G., Marcinko, J. P., Odell, L. H., Kirmeyer, G. J. and Wolfe, R. L. (1996). Occurrence of nitrification in chloraminated distribution systems. *Journal – American Water Works Association.* 88(7): 74-84.

64. Zeng, D. N., Fan, Z. Y., Chi, L., Wang, X., Qu, W. D and Quan, Z. X. (2013). Analysis of the bacterial communities associated with different drinking water treatment processes. *World Journal of Microbiology and Biotechnology*. 29: 1573-1584.

65. Zhang, Y., Oh, S. and Liu, W. T. (2017). Impact of drinking water treatment and distribution on the microbiome continuum: an ecological disturbance's perspective. *Environmental Microbiology*. 19(8): 3163-3174.

Figures

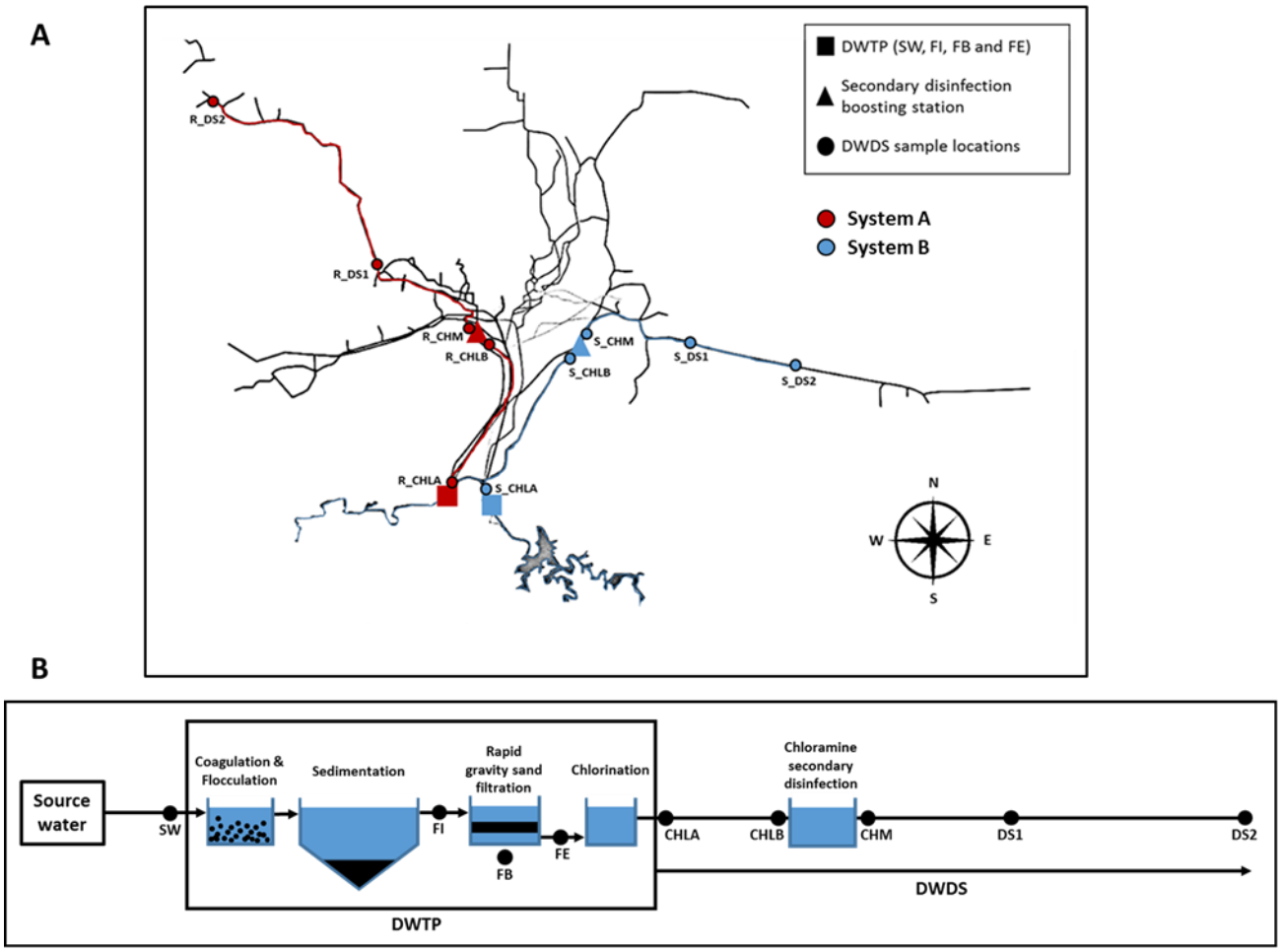


Figure 1: (A) Site map of the location of the drinking water treatment plants (R_DWTP and S_DWTP) and their corresponding distribution systems (R_DWDS and S_DWDS). System R is indicated in red and System S in blue. The two treatment plants are represented as squares, the two-secondary disinfection boosting stations, where chloramine is added, are represented as triangles and all sample locations are represented as circles. (B) Schematic of the layout of the DWTP and DWDS showing all sample locations. Within the two DWTPs source water (SW), filter inflow (FI), filter bed media (FB) and filter effluent (FE) samples were collected. All other sample locations are indicated on the figure and described in the text.

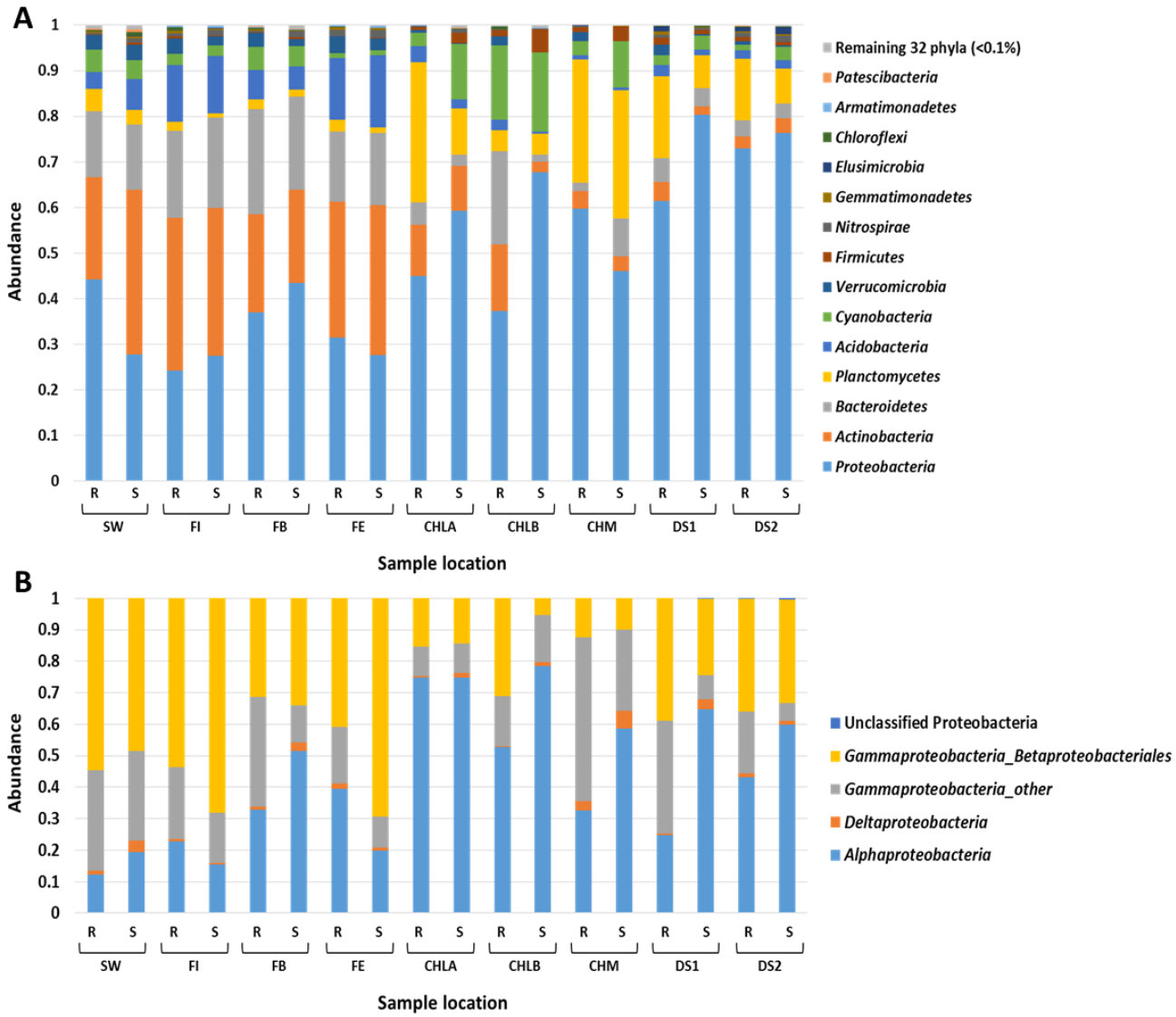


Figure 2: (A) Phylum-level mean relative abundance of bacterial sequences detected over the duration of the study at each sample location within the two DWTPs and corresponding DWDS (R and S DWDS sections). The 14 most abundant and unclassified phyla (> 0.1%) are shown here, with the remaining 32 phyla (< 0.1%) grouped together as a single group. Phyla are shown in the legend on the right of the figure. See Table S3 for mean relative abundances. (B) Mean relative abundance of proteobacterial classes detected over the duration of the study at each sample location for each system.

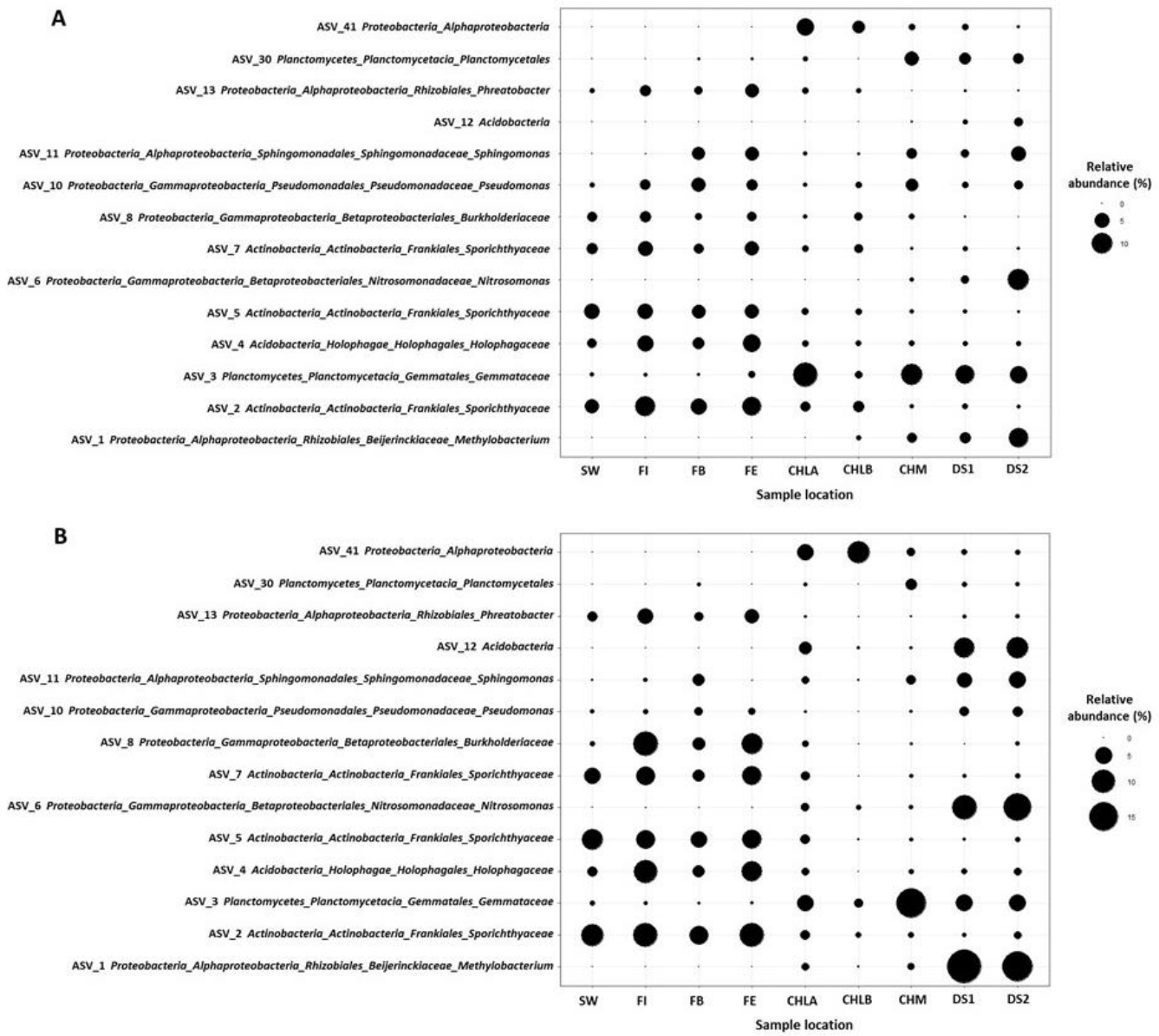
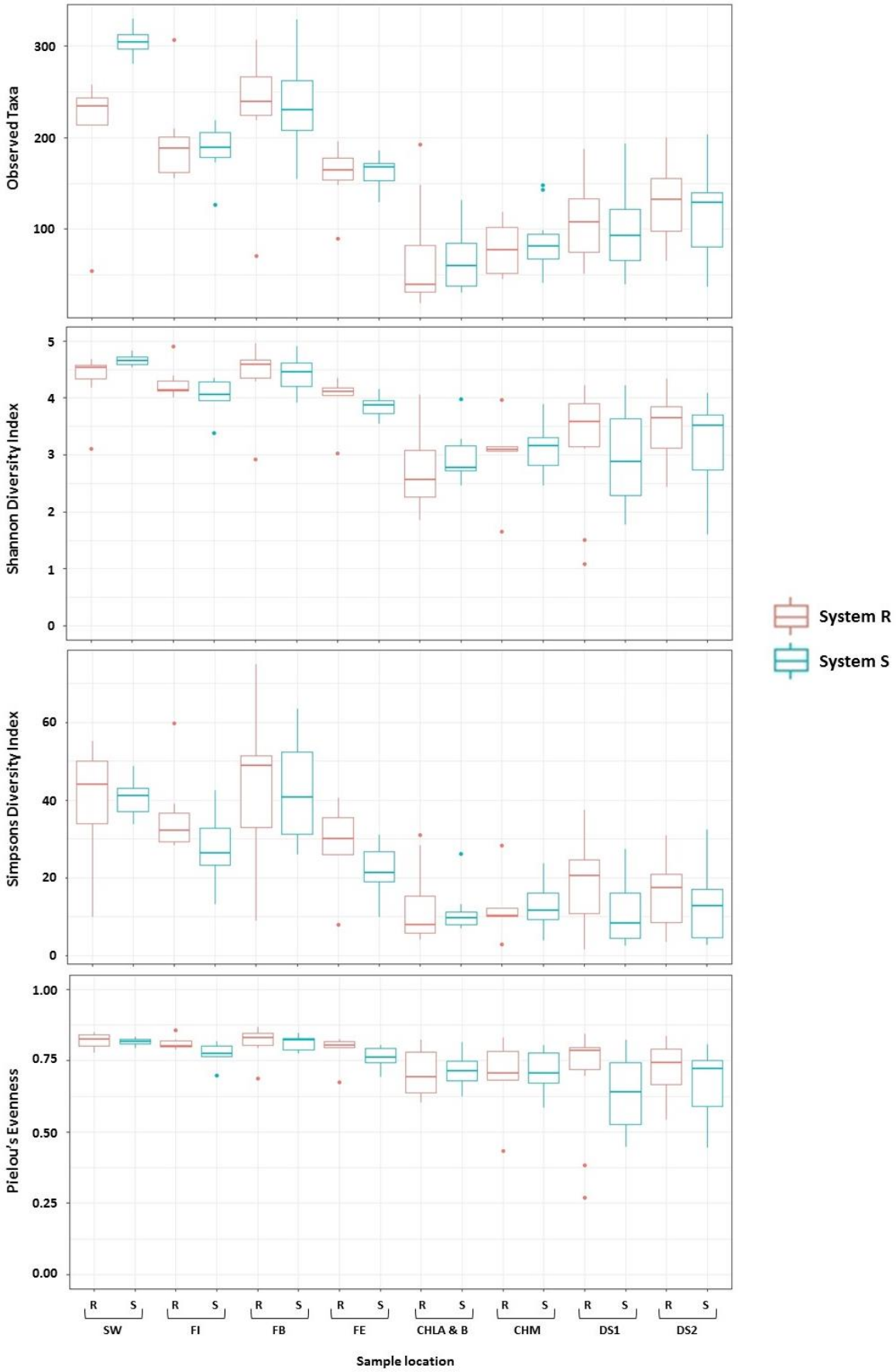
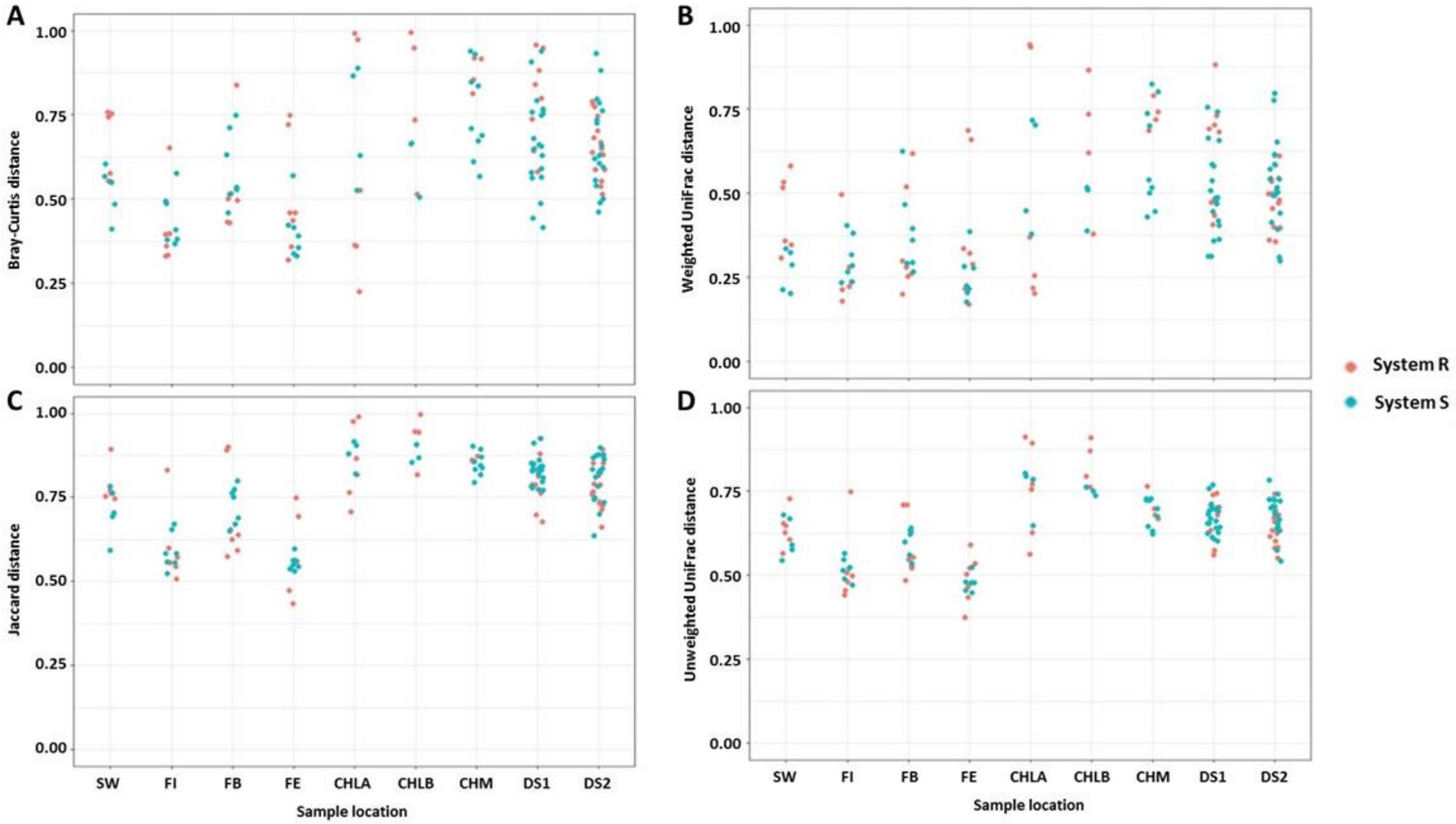


Figure 3: Variation in relative abundance of the 14 most abundant bacterial amplicon sequence variants (ASV's) with a mean relative abundance of ≥ 1% across all samples from (A) System R and (B) System S. The relative abundance for each sample location was averaged over duration of the study for each system. Percentage relative abundance of each ASV is indicated in the legends on the right if the figures. See Table S4A and S4B for mean relative abundances (MRA) of dominant ASVs.


Observed Taxa
300
200
100
Shannon Diversity Index
5
4
3
2
1
0
Simpsons Diversity Index
60
40
20
0
Pielou's Evenness
1.00
0.75
0.50
0.25
0.00
R S R S R S R S R S R S R S R S
SW FI FB FE CHLA & B CHM DS1 DS2
Sample location
System R
System S

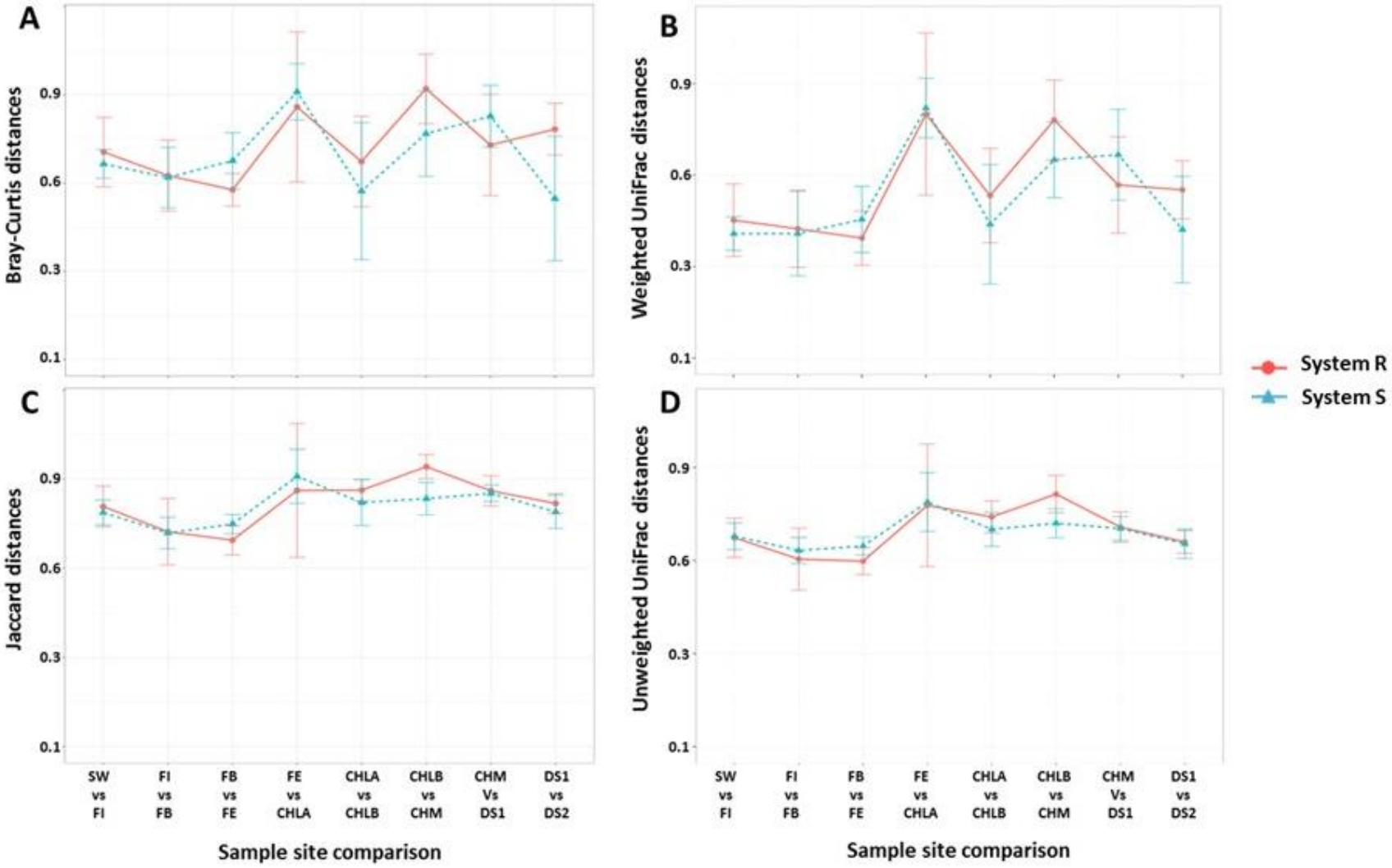
1196

Figure 4: Spatial changes in richness (observed taxa), diversity (Shannon Diversity Index and Inverse Simpson Diversity Index) and evenness (Pielou's evenness) averaged across all sampling locations for each month. Points represent all sample sites collected for each month. Samples coloured based on DWTP and corresponding DWDS (Lines R and S) (subsampled at 1263 iterations=1000).


A
B
C
D
Bray-Curtis distance
Weighted UniFrac distance
Jaccard distance
Unweighted UniFrac distance
1.00
0.75
0.50
0.25
0.00
SW
FI
FB
FE
CHLA
CHLB
CHM
DS1
DS2
Sample location
System R
System S

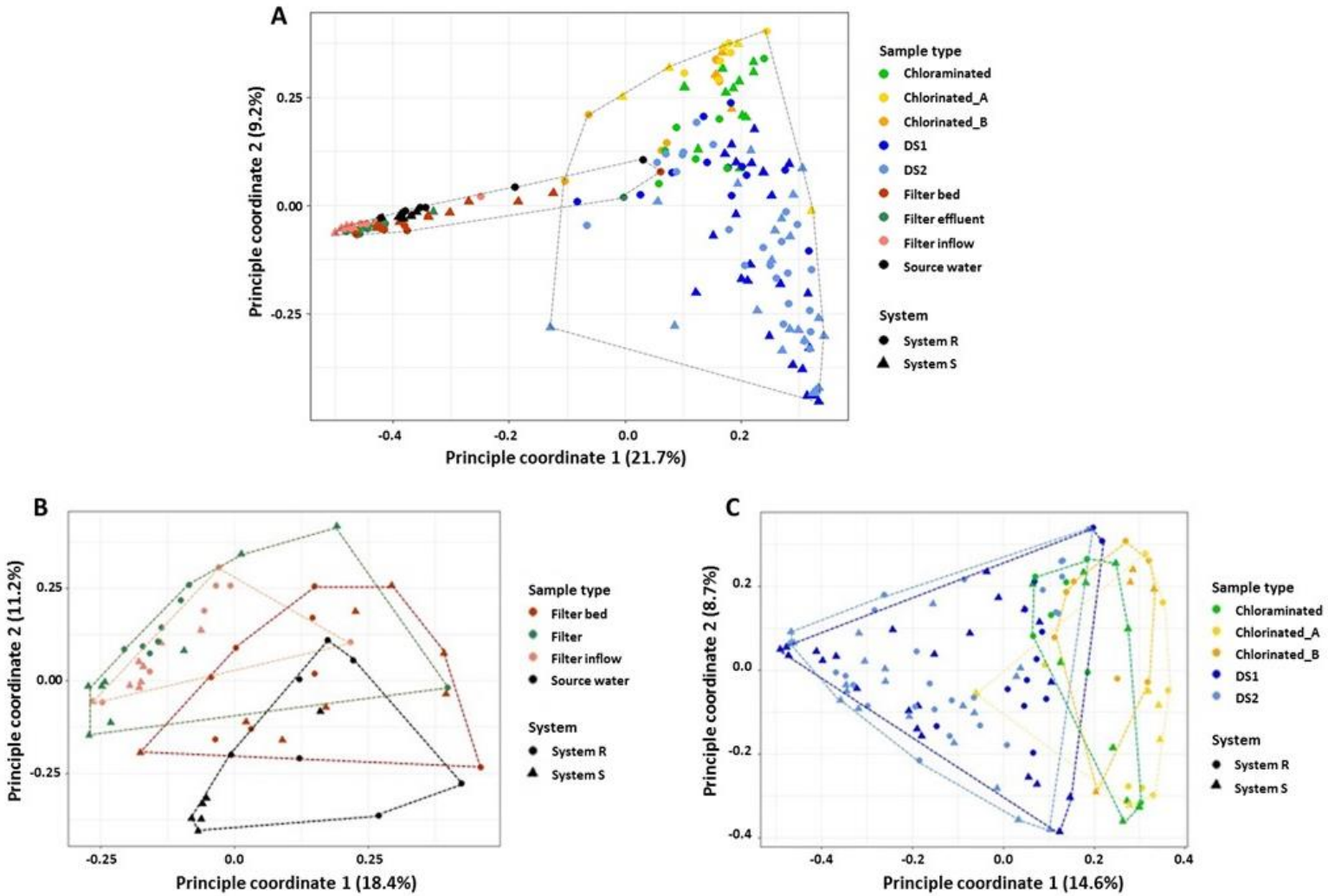
1202

Figure 5: Temporal variation within each sample location. Beta diversity pair-wise comparisons include samples from consecutive months within each location over the eight month study period for both structure based metrics: (A) Bray-Curtis, (B) Weighted UniFrac and membership based metrics: (C) Jaccard, (D) Unweighted UniFrac. Samples form System R are indicated in red and samples from System S are indicated in blue.


A
B
C
D
Bray-Curtis distances
Weighted UniFrac distances
Jaccard distances
Unweighted UniFrac distances
0.9
0.6
0.3
0.1
System R
System S
SW vs FI
FI vs FB
FB vs FE
FE vs CHLA
CHLA vs CHLB
CHLB vs CHM
CHM Vs DS1
DS1 vs DS2
Sample site comparison

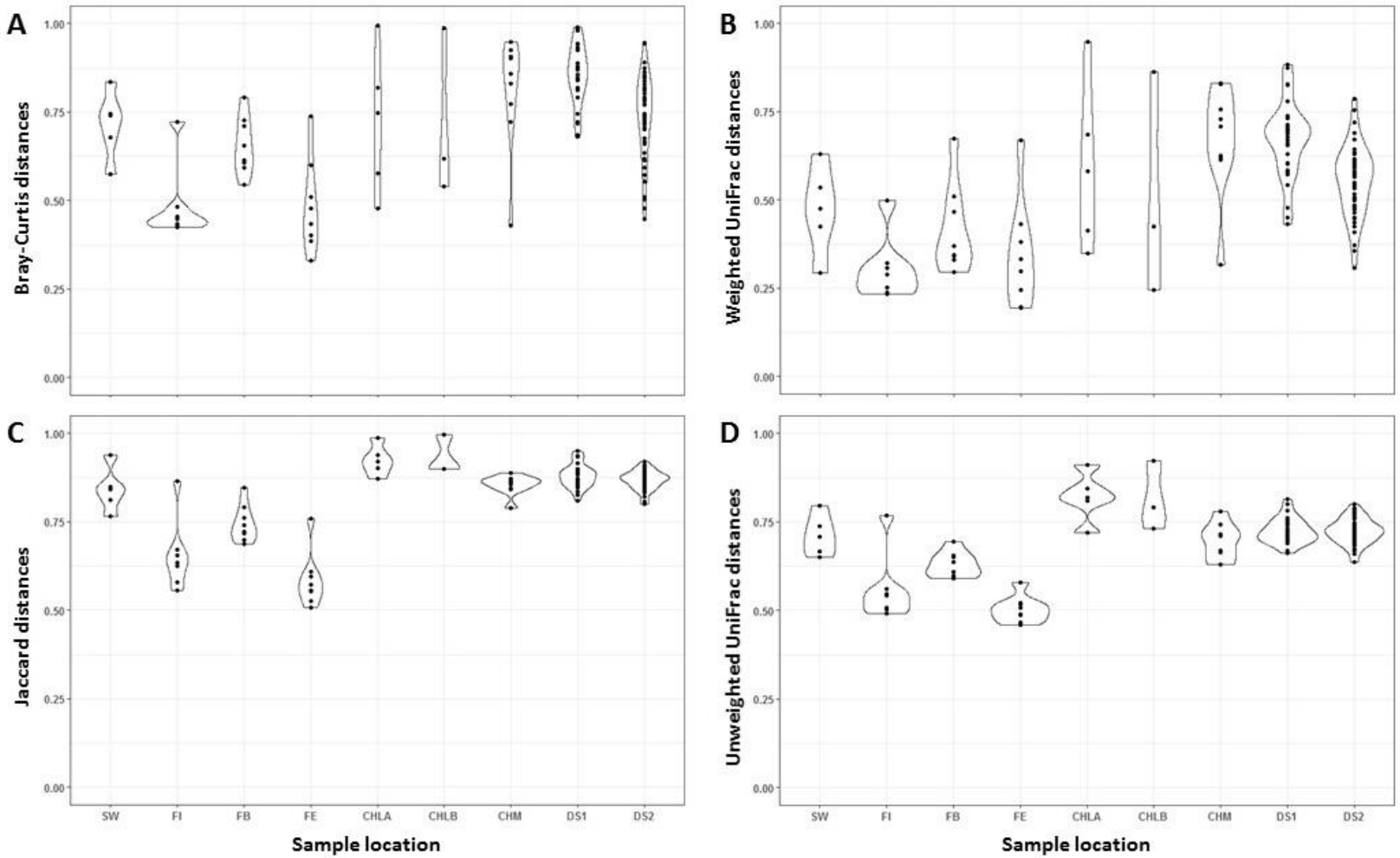
1209

Figure 6: Average pairwise beta diversity comparisons [structure based metrics: (A) Bray-Curtis, (B) Weighted UniFrac and membership based metrics: (C) Jaccard, (D) Unweighted UniFrac] between consecutive locations within each of the two systems for corresponding months. Sample comparisons from System R are indicated as red circles with a solid line and those from System S are indicated as blue triangles with a dashed line. Points indicate the mean and error bars indicate standard deviations.


A
Principle coordinate 2 (9.2%)
Principle coordinate 1 (21.7%)
Sample type
Chloraminated
Chlorinated_A
Chlorinated_B
DS1
DS2
Filter bed
Filter effluent
Filter inflow
Source water
System
System R
System S
B
Principle coordinate 2 (11.2%)
Principle coordinate 1 (18.4%)
Sample type
Filter bed
Filter
Filter inflow
Source water
System
System R
System S
C
Principle coordinate 2 (8.7%)
Principle coordinate 1 (14.6%)
Sample type
Chloraminated
Chlorinated_A
Chlorinated_B
DS1
DS2
System
System R
System S

1216

Figure 7: Principal coordinate analysis plot (based on Bray-Curtis dissimilarity) showing the spatial and temporal variability of the bacterial community structure among all samples from both systems (A), within the two DWTPs (B) and within the two corresponding DWDSs (C). Spatial groupings are shown where data points are coloured based on sample location and shaped based on the system they originate from (System R samples are indicated as circles and System S samples as triangles). Colour and shapes are indicated in the legends on the left of all plots.


A
B
C
D
Bray-Curtis distances
Weighted UniFrac distances
Jaccard distances
Unweighted UniFrac distances
1.00
0.75
0.50
0.25
0.00
SW
FI
FB
FE
CHLA
CHLB
CHM
DS1
DS2
Sample location
Sample location

1224

Figure 8: Pairwise beta diversity comparisons between corresponding sample locations from the two systems [structure based metrics: (A) Bray-Curtis, (B) Weighted UniFrac and membership based metrics: (C) Jaccard, (D) Unweighted UniFrac]. Sample abbreviations on the x-axis refer to source water (SW), filter inflow (FI), filter bed media (FB), filter effluent (FE), chlorinated water leaving the DWTP (CHLA), chlorinated water entering the secondary disinfection boosting station (CHLB), chloraminated water (CHM), distribution system site 1 (DS1) and distribution system site 2 (DS2). Pairwise beta diversity comparisons include samples from the same month. Mean and standard deviations of each comparison is shown in Table S6.