Microbial ecology and metabolism in chloraminated drinking water systems

By

Sarah Catherine Potgieter

Submitted in partial fulfilment of the requirements for the degree

Philosophiae Doctor (Microbiology)

In the

Department of Biochemistry, Genetics and Microbiology

Faculty of Natural and Agricultural Sciences

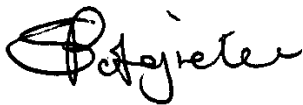University of Pretoria

Pretoria

April 2019

Supervisor: Prof S. N. Venter

Co Supervisor: Dr A. J. Pinto

DECLARATION

I, Sarah Catherine Potgieter declare that the thesis/dissertation, which I hereby submit for the degree *Philosophiae Doctor* (Microbiology) at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

01/04/2019

_____   _____

Sarah Catherine Potgieter   Date

Microbial ecology and metabolism in chloraminated drinking water systems

By

Sarah Catherine Potgieter

Supervisor: Prof S. N. Venter

Co Supervisor: Dr A. J. Pinto

Department: Biochemistry, Genetics and Microbiology

Degree: PhD (Microbiology)

SUMMARY

Drinking water systems are complex and multi-dimensional aquatic environments. The chemical, physical and biological factors inherent to each drinking water system drive changes in the microbial community dynamics and shape the drinking water microbiome. The complexity of interactions between these factors together with the use of differing source waters as well as the variability in treatment strategies and system design make each drinking water system unique. It is therefore important to understand the microbial ecology of drinking water systems and how the factors specific to each system shape the microbial community.

The aim of this project was to understand which factors drive the microbial community dynamics of a complex, large-scale South African drinking water system, with the application of multiple disinfectant regimes (i.e., chlorination, chloramination and hypochlorination). Specific objectives included: (i) investigating the long-term spatial and temporal microbial community dynamics along the distribution system, (ii) investigating the reproducibility of the microbial community dynamics of two drinking water systems treating similar source waters and lastly, (iii) in light of the use of chloramine as a secondary disinfectant and the observed dominance of *Nitrosomonas* ssp., investigate microbial mediated nitrogen metabolism in the chloraminated drinking water.

The large-scale drinking water system investigated in this project offered unique opportunities to explore the changes in microbial community dynamics across two drinking

water treatment plants, over large distribution distances as well as through different disinfection regimes. Initially, the study conducted over a two-year period, demonstrated that substantial temporal and seasonal trends existed, specifically at individual sample locations. However, when considering the distribution system as a whole, the spatial dynamics explained more of the variation in the microbial community. In addition, the microbial community dynamics were found to be reproducible, where treatment and distribution had a similar impact on the microbial community dynamics between two systems. Similar treatment operations resulted in the development of similar microbial communities. However, the microbial communities demonstrated a differential response to chlorination, but the selection of similar taxa through distribution indicated stabilisation of the microbial community post-disinfection.

Lastly, in light of the use of chloramine and based on the observed dominance of *Nitrosomonas* ssp., in the chloraminated sections of both the previous studies, a metagenomic approach was used to investigate the microbial mediated nitrogen metabolism. This revealed that *Nitrosomonas* and *Nitrospira* species dominated in chloraminated reservoirs and suggested that these taxa play significant roles in nitrification in chloraminated water. Furthermore, based on the dominant nitrogen transforming genes and metagenome assembled genomes, it was observed that the nitrate formed through nitrification is either reduced back to ammonia or to nitric oxide, where it may play a role in the regulation of biofilm formation.

These investigations highlighted the interplay between the spatial and temporal dynamics of a large-scale drinking water system and revealed the factors that drive the changes in the microbial community through treatment and distribution. In addition, this project provided insight into the genetic network behind microbially mediated nitrogen metabolism in chloraminated drinking water and may help improve the understanding of the processes behind nitrification in systems where it is a challenge. This project contributes to the current knowledge base in this field and provides drinking water utilities with the opportunity to understand the range of mechanisms that influence the microbial community and understand the underlying contributing factors that impact microbial growth in drinking water systems.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER 3 109

REPRODUCIBLE MICROBIAL COMMUNITY DYNAMICS OF TWO DRINKING WATER SYSTEMS TREATING SIMILAR SOURCE WATERS...............................109

LIST OF TABLES

CHAPTER 2

LIST OF FIGURES

CHAPTER 2

CHAPTER 4

LIST OF APPENDICES

Tables

1. Table A1: Samples excluded due to failed sequencing
2. Table A2: Taxonomy of the core community
3. Table A3: Phylum level classification
4. Table A4: Parameter data (Temperature and chlorine residuals)
5. Table A5: PERMANOVA results for all samples
6. Table A6: PERMANOVA results for section 1
7. Table A7: PERMANOVA results for section 2
8. Table A8: PERMANOVA results for section 3

Figures

1. Figure A1: Temporal changes in alpha diversity measures
2. Figure A2: Principal-coordinate (PCoA) plots
3. Figure A3: Temporal variation in bacterial community structure and membership
4. Figure A4: Spatial changes in richness
5. Figure A5: Pair wise beta-diversity distances between DWDS sample locations

Tables

1. Table B1**:** Water quality parameters of the source water from both systems
2. Table B2A**:** Water quality parameters for DWDS samples from System R
3. Table B2B**:** Water quality parameters for DWDS samples from System S
4. Table B3: Mean relative abundance of dominant bacterial phyla
5. Table B4A: Mean relative abundance of the most abundant ASVs in System R
6. Table B4B: Mean relative abundance of the most abundant ASVs in System S
7. Table B5: Alpha diversity measures for each study site
8. Table B6: Beta diversity comparisons

Figures

1. Figure B1: Mean relative abundance of the 14 most abundant bacterial ASV's
2. Figure B2: Venn diagram showing the shared ASV's between the two source waters.

LIST OF ABBREVIATIONS

AAI	Average amino acid identity
ANOVA	Analysis of Variance
AMOVA	Analysis of Molecular Variance
AOA	Ammonia oxidising archaea
AOB	Ammonia oxidising bacteria
AOC	Assimilable organic carbon
AOM	Ammonia oxidising microorganisms
ATP	Adenosine triphosphate
BOM	Biodegradable organic matter
BDOC	Biodegradable dissolved organic carbon
CHL	Chlorinated water
CHLA	Chlorinated water leaving the treatment plant
CHLB	Chlorinated water entering the booster station before chloramination
CHM	Chloraminated water
DADA2	Divisive Amplicon Denoising Algorithm
DBP	Disinfection by-products
DGGE	Denaturing gradient gel electrophoresis
DNA	Deoxyribonucleic acid
DOC	Dissolved organic carbon
DS	Distribution system sample points
DWDS(s)	Drinking water distribution system(s)
DWTP(s)	Drinking water treatment plant
EPS	Extracellular polymeric substances
FB	Filter bed media
FCM	Flow cytometry
FE	Filter effluent
FI	Filter influent
FST	F-statistic
GAC	Granulated activated carbon
HAA	Haloacetic acids

HAN	Haloacetonitriles
HCHL	Hypochlorinated water
HNM	Halonitromethanes
HPC	Heterotrophic plate counts
HSD	Tukey Honest Significant Differences
M	Mean
MAG(s)	Metagenome assembled genome(s)
MIC	Microbial mediated corrosion
MRA	Mean relative abundance
NGS	Next generation sequencing
NOB	Nitrite oxidising bacteria
NOM	Natural organic matter
ORF	Open reading frame
OTU	Operational taxonomic unit
PCoA	Principal coordinate analysis
PCR	Polymerase chain reaction
PERMANOVA	Permutational analysis of variance
PI	Propidium iodide
PNR	Potential nitrification rate
PoU	Point of use
RES	Reservoir
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
RSF	Rapid sand filters
SD	Standard deviation
SEQ(s)	Sequence variants
SSCP	Single-strand conformation polymorphism
SSF	Slow sand filters
SSU	Small subunit
SW	Source water
THM	Trihalomethanes
TOC	Total organic carbon

TCC	Total cell counts
T-RFLP	Terminal restriction fragment length polymorphism

PREFACE

Microbial ecology is a rapidly growing field with the development of high throughput sequencing providing new insights into the composition, structure, functions and dynamics of microbial communities. Microbial communities can be extremely diverse with only a fraction of microorganism in natural environments being culturable. With this in mind, recent studies involving the drinking water environment have revealed that drinking water treatment and distribution systems are complex and variable environments for microbial growth. The high level of variability in physical, chemical and biological characteristics inherent to each drinking water system results in a taxonomically diverse drinking water microbiome that ultimately drives critical processes that directly impact water quality.

Information on the microbial community composition, function dynamics of established South African drinking water systems is sorely lacking. Currently, approximately 5 million people in South Africa do not have access to potable water. Many small water works struggle to provide adequate treatment and disinfection of the water, resulting in increased incidences of waterborne diseases. In light of this, the majority of studies performed on South African water systems have concentrated on water systems in rural communities, the presence of waterborne pathogens and their associated health risks.

Therefore, this project offers an opportunity to study a large-scale South African drinking water system, including two treatment plants and their corresponding distribution systems. Here, source water is primarily drawn from a river and dam system and treated via two treatment plants, which perform the same consecutive treatment strategies. The distribution network consists of large diameter pipelines stretching over 3056 kilometers in its entirety. The distribution system feeds 58 strategically located service reservoirs, which supplies large metropolitan and local municipalities including approximately 12 million people, as well as mines and industries with on average, 4800 million litres of water daily (Rand Water). In light of this, questions arise relating to the microbial community dynamics of such a large drinking water treatment and distribution system. In addition, this drinking water system employs a combination of disinfection treatments (chlorination and chloramination), raising questions regarding the effect of the combination of disinfection strategies on the microbial community ecology and dynamics. Furthermore, due to the use of chloramine as a secondary

disinfectant in this system and together with the observed increased potential of nitrification, the metabolic fate of the ammonia added through chloramination needs to be investigated.

Therefore, this project aims to: (i) understand the long-term spatial and temporal dynamics of the microbial community within this large-scale drinking water distribution system using multiple disinfectant regimes, (ii) investigate the reproducibility of the microbial community dynamics of the two drinking water systems treating similar source water with the same treatment strategies and distribution, and (iii) in light of nitrification, explore nitrogen metabolism in chloraminated drinking water systems.

By contributing to the current knowledge base in this field, this study provides the opportunity for drinking water utilities to understand the range of mechanisms that influence the microbial community dynamics, over varying temporal scales and/or operational stages. This may help water utilities better understand the impact of treatment and the downstream dynamics of diverging sections within large-scale distribution systems on the drinking water microbiome. And lastly, this study provides insight into the genetic network behind microbially mediated nitrogen metabolism in chloraminated drinking water and aims to improve control strategies when nitrification becomes an operational problem. Ultimately, understanding the microbial community dynamics in a South African DWDSs can lead to the long-term improvement of drinking water management and operation processes.

CHAPTER 1

THE MICROBIAL ECOLOGY OF DRINKING WATER SYSTEMS: A LITERATURE REVIEW

Potgieter, S. C., Pinto, A. J. and Venter S. N. The microbial ecology of drinking water systems: A review. (In preparation)

Author Contributions:

SC Potgieter and SN Venter were involved in the conceptualisation and layout of review. All authors contributed to the concept and drafting of the review.

1. Introduction into microbial ecology of drinking water

Drinking water treatment and distribution systems are recognised as complex aquatic environments offering multiple unique habitats that support microbial growth through the different stages of treatment and distribution. Drinking water leaving the treatment plant may be high quality drinking water. However, treated water is subjected to various conditions within the distribution system, which ultimately leads to the deterioration of the water quality. Therefore, maintaining the biologically stability of drinking water through distribution is one of the biggest obstacles facing drinking water utilities.

Biologically stable drinking water is typically defined as water that does not support the growth of microorganisms to a sugnificant extent (Rittmann and Snoeyink, 1984; Prest *et al.*, 2016). Therefore, biological stability of drinking water can only be maintained based on a clear understanding of the microbial composition and ecology within drinking water systems. The microbial ecology of drinking water systems is governed by multiple environmental and engineering factors as well as operational conditions that influence the composition and structure of microbial communities present in the bulk water, biofilms, and sediments (Wang *et al.*, 2014; El-Chakhtoura *et al.*, 2015). Through the study of microbial ecology, the mechanisms that govern the assembly of microbial communities and the factors that drive changes in community composition and structure, along different spatial and temporal scales, can be determined (Hanson *et al.*, 2012; Nemergut *et al.*, 2013). Understanding the microbial ecology will then help answer questions regarding the community's origin and stability as well as the interactions that restrict them or keep them stable, i.e. the interactions between members of the community and their environment (Gülay *et al.*, 2016).

Microbial communities are not limited to a single taxonomic group but are usually highly diverse. Typically, the majority of the observed membership within the community are rare taxa, present at low relative abundances, with only a small portion of taxa occurring at high abundances. These low-abundant taxa may increase the diversity of the community and have a significant impact on the dynamics of the community as a whole (Shade *et al.*, 2014). They have also been shown to be important indicators of environmental disturbances over time (Szabó *et al.*, 2007; Gülay *et al.*, 2016).

In microbial ecology, diversity is typically defined in terms of alpha and beta diversity. The inventory or alpha diversity, describes the diversity found within an environment on the smallest spatial scale (within a sampling point). In contrast, beta diversity, also referred to as differentiation diversity, measures the diversity between environments (amongst/between sampling points) (Nemergut *et al.*, 2013). Both alpha and beta diversity are defined by species richness and species abundance information using different diversity metrics and indices (Little *et al.*, 2008).

Many alpha and beta diversity studies have focussed on the highly abundant species, constituting the core microbial community, which are often linked to microbial-mediated processes (Albers *et al.*, 2015; Gülay *et al.*, 2016). For example, microbial communities within drinking water sand filtration systems may play a beneficial role through biofiltration. Alternatively, microbial growth in drinking water distribution systems (DWDS) can negatively impact water quality through the presence of potential pathogens, contributing to microbially induced corrosion of the DWDS infrastructure, disinfectants decay and nitrification (Pinto *et al.*, 2012a; Camper, 2014; Roeselers *et al.*, 2015; Douterelo *et al.*, 2018a). Therefore, some of the major challenges facing water utilities is to minimise these negative aspects and supply microbially safe drinking water to communities through limiting microbial growth and maintaining microbially stable drinking water throughout distribution.

Supplying drinking water that is both chemically and microbially safe as well as acceptable in terms of taste, odour and appearance is vital to public health and economic growth (Geldreich, 1996; Hunter *et al.,* 2010). One of the goals of microbial ecology is to determine the factors that drive and maintain the structure and composition of microbial communities. Therefore, understanding the microbial ecology of drinking water microbial communities can improve the knowledge of their interactions with the environment and each other. With this growing body of knowledge, this will potentially aid in the development of more effective microbial control strategies for drinking water industries in the long-term (Bautista-de los Santos *et al.*, 2016).

2. Current understanding of drinking water distribution systems

2.1 System design: treatment and distribution

Drinking water is one of the most highly monitored and regulated resources. The delivery of safe water and sufficient water supply depends largely on the reliability of the infrastructure as well as the availability and quality of the source water (Bondank *et al.*, 2018). To produce microbially safe drinking water, various treatment processes are applied to appropriately treat source waters (Hwang *et al.*, 2012b). Regardless of these efforts, microbial communities persist in the distribution system. Microbial communities in drinking water migrate through the drinking water treatment plant (DWTP), into the drinking water distribution system (DWDS) and into the built environment and consumer homes (Pinto *et al.*, 2014). DWDS are designed and built to prevent intrusion and distribute drinking water to the consumer. The movement of microbial communities through the system is an integral component of drinking water dynamics and maintaining and managing the distribution system is a highly complicated process but vital in the delivery of safe drinking water (Bautista-de los Santos, *et al.*, 2016).

There are many factors that could impact the water quality as well as the microbial community composition and structure (Liao *et al.*, 2015a; Prest *et al.*, 2016a). An important factor is the choice of treatment strategy, where the treatment regime is usually highly site specific as it is selected based on the characteristics of the source water, which can be highly diverse. Groundwater typically contains low organic nutrients as well as lower bacterial numbers and may require minimal treatment (Prest *et al.*, 2016a). Alternatively, surface waters contain high levels of organics and bacterial cell numbers and may require additional treatment steps. The selected treatment regime determines various quality aspects of the finished treated water such as the concentration of organic and inorganic nutrients as well as the microbial community composition and abundance within the DWDS (Prest *et al.*, 2016a; Li *et al.*, 2017).

In the case of surface waters with a high levels of natural organic matter (NOM), processes such as coagulation, flocculation and sedimentation may be applied (Edzwald, 1993; Matilainen *et al.*, 2010). These initial treatments may involve the removal of large particles and decreasing turbidity depending on the composition of the source water. Typically sand

filtration is used to decrease turbidity and remove bacterial growth substrates to limit bacterial growth downstream in the distribution system (Pinto *et al*., 2012a; Liao *et al*., 2015b). Biological filtration steps currently used include, rapid sand filters (RSF), granulated activated carbon (GAC) and/or slow sand filters (SSF) (Wang *et al*., 2013; Lautenschlager *et al*., 2014). However, it has been shown that the microbial community present in the sand filter can seed and shape the water microbiome downstream (Pinto *et al*., 2012a; Bruno *et al*., 2018). Bacterial growth within the distribution system is limited and bacteria are inactivated through the application of disinfection processes including, ozonation, the addition of chlorine and chloramination and UV treatment (Norton and LeChevallier, 1997; von Gunten, 2003; Hwang *et al*., 2012b; Camper, 2014).

However, following all attempts to inactivate bacteria and limit microbial growth through treatment processes, microbial communities still persist under extreme conditions of low nutrient concentrations and disinfectant residuals (Pinto *et al.,* 2014; Bertelli *et al*., 2018). After leaving the DWTP and with increased contact with the DWDS infrastructure the water quality typically deteriorates, both hygienically and aesthetically (Boe-Hansen *et al.,* 2002). In the DWDS, water is moved between reservoirs, through various pipelines (differing in composition and age) extending across vast areas, and eventually through the plumbing systems of consumer homes (Douterelo *et al*., 2017). Plumbing systems in buildings and homes may be a significant source of contamination as these systems introduce possible dead ends and a variety of attachment sites for microbial growth (Flemming *et al*., 2014). Drinking water leaving the consumer tap may contain up to $10^6 – 10^8$ microbial cells per litre (Hammes *et al*., 2008).

2.2 Microbial growth in drinking water systems

Water distribution systems may be viewed as multifaceted aquatic environments with highly diverse microbial communities, including bacteria, archaea, eukaryota and viruses (You *et al*., 2009; Thomas and Ashbolt, 2011; Siqueira and Lima, 2013; Liu *et al.,* 2013a; Pinto *et al*., 2014; Gall *et al*., 2015). The treated water enters the distribution system with a physical particle load, a nutrient load as well as a microbial load (although reduced compared to the source water) (REFS). Microorganisms within DWDSs may originate from the source water (Gomez-Alvarez *et al*., 2015), survive the initial treatment processes, after which they enter the distribution system. These microbial populations are adapted to survive and grow within

the oligotrophic environment of DWDSs (Douterelo *et al*., 2018a). The majority of these microorganisms are inherent manageble system components that grow normally in drinking water systems. Alternatively, microorganisms can potentially enter the system through external sources such as open reservoirs, pipe breakages, maintenance and the reduction of flow velocity resulting in back washing.

Liu *et al.,* (2013a and 2014) describes the DWDS as being multi-dimensional, where there are typically four phases within the system that serve as available microenvironments for microbial growth. These four phases are defined and summarised as the pipe wall biofilm (formed on the inner surface of the pipe material), the bulk water (the water phase flowing through the pipe), suspended solids (the particulate matter passing through the system) and loose deposits (particulate matter settled on the bottom of the pipe) (Liu *et al*., 2014). There are few studies focused on the contribution of suspended solids and loose deposits but it has been suggested that they should not be overlooked, as microbial growth in loose deposits may be comparable to that of the biofilms lining the pipe walls (Liu *et al.,* 2013a and 2013b).

Bacteria dominating DWDS are oligotrophic as they survive in environments with low substrate concentrations. In short, the DWDS microbiome is dominated typically by the phyla *Firmicutes* and *Proteobacteria,* specifically the proteobacterial classes *Alphaproteobacteria, Betaproteobacteria* (recently reclassified as order *Betaproteobacteriales*, phylum *Gammaproteobacteria*) and *Gammaproteobacteria* (Hong *et al.,* 2010; Kwon *et al.,* 2011; Douterelo *et al.,* 2014; Sun *et al.,* 2014; Parks *et al*., 2018). The proportion of which these bacterial groups exist within drinking water systems is highly dependent on multiple factors such as source water, disinfectant strategy (Hwang *et al.,* 2012b), filtration processes (Kwon *et al.,* 2011; Pinto *et al.,* 2012a) and pipe materials. Hwang *et al.* (2012b) describes a microbial community study where changes in community composition were directly related to changes in disinfectant type (Chlorination and chloramination). Here they revealed that *Cyanobacteria, Methylobacteriaceae, Sphingomonadaceae* and *Xanthomonadaceae* were more abundant in chlorinated water, whereas *Methylophilaceae, Methylococcaceae* and *Pseudomonadaceae* were abundant in chloraminated water. Furthermore, nitrifying bacteria and archaea have been found to be more abundant in chloraminated systems (Wolfe *et al.,* 1990; Cunliffe, 1991; Regan *et al*., 2003; Sathasivan *et al*., 2008).

2.2.1   The influence of biofilm development on inner pipe surfaces

Biofilms are present in all DWDS, where the interface between water and the pipe wall serves as a site for the build-up of organic matter and cells, leading to bacterial attachment and multiplication (Batté *et al*., 2003). A biofilm is typically described as a layer of microorganisms connected by extracellular polymeric substances (EPS), which facilitates attachment to pipe surfaces as well as protection from disinfectants (Wingender and Flemming, 2011). Typically, bacteria with the abillity for biofilm formation are able to produce increased quantities of EPS, thereby allowing them to initiate biofilm formation and expand the matix (Douterelo *et al*., 2013).

As the bulk water moves through the distribution system, biofilms develop on the inner surfaces of the pipelines (Liu *et al.,* 2013a). Their distribution over the pipe surface is non-uniform and their structure is discontinuous and heterogeneous. In addition, they consist of a mixture of different microorganisms with various activities depending on their position in the aggregate (Batte *et al*., 2003). The biofilm is potentially always reorganising as detachment, hydraulic forces, levels of nutrients and introduction of new members are constantly shaping their development (Batté *et al*., 2003). With this in mind, a biofilm may never reach a steady state as selection is always occurring and changes in the environmental conditions may favour different bacterial species (Boe-Hansen *et al.,* 2002). The bacterial activity also differs depending on the biofilm age, where the growth rate of a mature biofilm may be very different to the rate during initial colonisation (Boe-Hansen *et al.,* 2002).

Biofilms become important when they negatively affect DWDSs. Biofilms promote the deterioration of pipe surfaces, leading to microbial mediated corrosion (MIC), the growth of ammonia and nitrite oxidisers contribute to nitrification in chloraminated systems (Wilczak *et al.,* 1996) and they may harbour coliforms and potential pathogens (Batte *et al*., 2003; Ndiongue *et al.,* 2005). In addition, bacteria and fungi in established biofilms display increased resistance to disinfectant residuals and can withstand higher disinfection concentrations than their planktonic counterparts in the bulk water (Berry *et al.,* 2006). As a result of this potential resistance, biofilms increase the disinfectant demand and promote disinfectant decay (Fish and Boxall, 2018). The presence of chlorine may promote the formation of biofilms as high chlorine concentrations can result in microorganisms favouring a biofilm state as a potential defence mechanism (Srinivasan *et al*., 2008; Fish and Boxall,

2018). Therefore, the control of biofilm development in DWDS requires a combination of actions including limiting the availability of nutrients such as biodegradable organic matter (BOM), mechanical cleaning and flushing and disinfection (Batte *et al*., 2003; Fish and Boxall, 2018)

Biofilms represent active zones for biomass accumulation and are the main form of organic material and microbial loading in DWDS. The proliferation of bacteria within the bulk water was thought to originate from biofilm detachment or shearing from the pipe wall (Batte *et al*., 2003; Wingender and Flemming, 2011; Fish and Boxall, 2018). It was previously believed that microbial growth within DWDSs originates primarily from biofilm growth on the internal surface of the drinking water pipelines with only a small portion of the microbial community existing freely in the water phase (Flemming *et al.,* 2002; Farkas *et al.,* 2013). However, the microbial communities between biofilms and bulk water have also been shown to be distinct in both community composition and gene function (Norton and LeChevallier, 2000; Henne *et al*., 2012; Douterelo *et al*., 2013; Liu *et al.,* 2014; Douterelo *et al*., 2018b). A study by LeChevallier *et al.* (1988) showed that the bacteria in the bulk water were different to those attached to the pipe surface and therefore had a minor impact on the biofilms. Studies from both Henne *et al*. (2012) and Liu *et al*. (2014) observed differences in the bacterial community composition between bulk water and biofilms, where different core bacterial communities were observed for both biofilms and bulk water. More recently, using a shotgun metagenomic approach, Douterelo *et al*. (2018b) were able to show differences in gene functions between biofilm and bulk water communities. Here they identified several resistance mechanisms in the biofilm community, including genes involved in biofilm formation and damage repair to external stressors such as chlorine and antibiotics.

2.2.2   Microbial growth in the bulk water

The bulk water phase is considered as the medium for the spread of microorganisms, nutrients and particles through the distribution system (Liu *et al.,* 2013a). It has long been assumed that bacteria in the bulk water originate from detachment of biofilms or re-suspension of the sediments rather than bacterial growth in the bulk phase itself (Prest *et al*., 2016a). In 1989, van der Wende *et al.* stated that detachment of the biofilm was responsible for the planktonic cells present in the bulk water and that bacterial growth in the bulk water was negligible.

However, biofilms and loose deposits may not dominate the distribution system under all conditions as previously thought. Boe-Hansen *et al.* (2002) showed using ATP content and leucine incorporation, that bulk water bacteria had a higher growth rate than biofilm bacteria. They suggested that the growth of bulk water bacteria should not be over looked as their growth is significant and should be considered in bacterial growth models in DWDS. Srinivasan *et al.* (2008) suggested that in parts of the distribution system, where chlorine residuals were low, bulk water bacterial may dominate. In addition, Douterelo *et al.* (2013) observed that despite having the same origin, bulk water and biofilms showed differences in the microbial community composition and structure, where bacterial groups in the bulk water are known to be more resistant to chlorine.

Furthermore, Pinto *et al.* (2012a) showed that the composition of the bulk water community was consistently shaped by the filter bed and the dominant bacteria associated with it. With this in mind and considering the slow growth rate of oligotrophic bacteria, the microbial community within the filter may be more important for bulk water bacteria. Pinto *et al.* (2012a) therefore suggested that the bacterial community on the filter could be used to predict the bacterial communities downstream in the distribution system.

These contradicting results may be a result of the variability in different distribution systems and the multiple factors that affect the microbial growth in both the biofilm and bulk water. The interaction between bacterial communities between biofilms, loose deposits and the bulk water is unclear. When considering the effects of hydraulic forces, the detachment of biofilms and re-suspension of sediments can all undoubtedly contribute to bacterial community composition and cell concentrations in the bulk water (Prest *et al.*, 2016a). These findings suggests that the bulk water potentially serves as a seed bank for sediments and biofilms and thereafter each phase develops its own bacterial community depending on the site specific environmental conditions (Henne *et al.*, 2012; Liu *et al.*, 2014; Prest *et al.*, 2016a).

2.2.3 Microbial growth associated with loose deposits

Several studies have discovered that loose deposits can potentially play a significant role in contributing to the microbial community within DWDSs (Liu *et al.*, 2014). Under favourable hydraulic conditions, the accumulation of loose deposits on the pipe bottom potentially serve as an attachment site for bacteria and source of organic compounds resulting in a

microenvironment for microbial growth (Prest *et al*., 2016a). In addition, these loose deposits allow for protection of bacteria from disinfectants and due to their mobility within the system, the associated bacteria can easily reach the consumer tap (Liu *et al*., 2014).

The bacterial communities associated with loose deposits are both highly variable and diverse in comparison to other phases. The community composition is occasionally similar to that of suspended solids, suggesting that loose deposits may simply be a result of sedimentation of the suspended solids. The community included members that were both aerobic and anaerobic with many of the bacterial genera identified as those involved in sulphur and nitrogen as well as iron and arsenic biogeochemical cycling (Liu *et al*., 2014). These findings correspond to the composition of elements typically found in loose deposits and thereby contribute to increased corrosion processes (Sun *et al*., 2014; Prest *et al*., 2016a).

There is a complex relationship between the bacterial communities within the biofilm, loose deposits and the bulk water, where the majority of research has concentrated on the biofilms lining the pipe walls and the more easily accessible bulk water. It was initially estimated that 95% of the microbial biomass in the DWDS is located in the biofilms surrounding the pipeline surfaces while the remaining 5% reside in the bulk water phase (Flemming *et al.,* 2002, Liu *et al.* 2013a and 2014). However, loose deposits and sediments have been overlooked due to difficulties in sampling (Liu *et al*., 2013a). A study by Liu *et al*. (2014) showed that together loose deposits and biofilms contribute 98% of bacterial cells where 60 – 90% were situated in the sediment phase. However, their impact on water quality is unclear.

2.3 Factors affecting the growth of the drinking water microbiome

DWDSs have highly variable physico-chemical properties inherent to each system (Douterelo *et al*., 2018a). It is therefore important to understand the factors influencing microbial growth and the interactions of these factors with microbial processes in order to facilitate effective control of microbial growth in drinking water systems. Conditions within the DWDS can significantly influence the biological stability of the drinking water. These factors include concentration of nutrients and organic compounds, disinfectant type, water temperature, pipe material and distribution system infrastructure, hydraulic forces, and as well as residence times. Often correlations between microbial growth and a single factor are not possible as

typically it is a combination of multiple factors and their interactions that affect microbial growth (Camper, 2014).

2.3.1   Concentration of nutrients and organic matter

Heterotrophic bacteria require organic carbon for growth, therefore the presence of organic matter can be responsible for the proliferation of biofilms and microbial growth in DWDSs (Camper, 2014). The availability of nutrients in the form of organic matter, nitrogen and phosphorus can all influence bacterial regrowth and biofilm formation as well as promote disinfectant decay (Chandy and Angles, 2001; Chu *et al.,* 2005). Nutrients mainly enter the distribution system in the form of biodegradable organic matter (BOM) and by limiting the BOM microbial growth can be controlled (Volk and LeChevallier, 1999; Chandy and Angles, 2001). Bacterial growth essentially requires carbon, nitrogen and phosphorus in the molar ratio of approximately 100C: 10N: 1P. As carbon is required the most, it is often considered as limiting in drinking waters. In many Eurpoean countries with disinfectant free systems, one of the key strategies to control microbial regrowth is to remove/limit nutrients during drinking water treatment (Chandy and Angles, 2001).

Biodegradable compounds typically either originate from the source water or from materials in contact with the bulk water, such as the pipe surfaces. These organic compounds act as carbon sources for heterotrophic bacteria, where oxygen is used as a hydrogen acceptor. Approximately 50% of the total organic carbon (TOC) is used in respiration and is released as CO_2, where the remaining 50% is assimilated into cellular components, contributing to cell growth (Momba *et al.,* 2000). BOM consists of a broad spectrum of different organic carbon compounds including simple organic sugars and acids as well as complex polymeric substances, such as humic compounds (Prest *et al.*, 2016a). BOM can typically be broken down into two main components: dissolved organic carbon (DOC) and assimilable organic carbon (AOC). Only a small portion of DOC can be utilised by bacteria as an energy source (Prest *et al.*, 2016a). The DOC is a broad term for organic compounds from a wide variety of sources and variable compositions, a fraction of which is biodegradable dissolved organic carbon (BDOC) (Servais *et al.*, 1989). BDOC can be defined as the portion of organic carbon metabolised and mineralised by heterotrophic bacteria (Camper *et al.*, 2014). The AOC can be described as the fraction of the BOM that can be easily used and converted into cell mass and can therefore give an indication of the growth potential of the heterotrophic bacteria present in the system (Van der Kooij, *et al.*, 1982; Van der Kooij, 1992; Camper *et al.*, 2014).

While many studies have focused on organic carbon and its influence on microbial growth, inorganic compounds such as nitrogen, phosphorous and trace elements (iron, potassium, copper, magnesium etc.) also contribute to the growth of heterotrophic bacteria, although to a lesser extent (Chandy and Angles, 2001; Prest *et al*., 2016a; Nescerecka *et al*., 2018). It has been shown that by limiting the availability of phosphorus, microbial growth can be controlled. Meittinen *et al,* (1997), showed that microbial growth is highly regulated by phosphorus as well as inorganic carbon and that phosphorus was the only inorganic element that had an impact on microbial growth. These results were confirmed by both Lehtola *et al,* (2002) and Nescerecka *et al*. (2018), where microbial growth was limited when microbially available phosphorus (MAP) was low.

Furthermore, the type and concentration of organic and inorganic substrates determine the type of microorganisms present in the water (Prest *et al*., 2016a). Though heterotrophic bacteria are known to dominate, the presence of autotrophic bacteria such ammonia and nitrite-oxidisers, iron-oxidisers, sulphate-reducing etc. have been observed in different drinking water environments. Examples of these include higher abundances of ammonia-oxidising bacteria such as *Nitrospira* and *Nitrosomonas* in waters rich in ammonium or chloraminated waters (Wolfe *et al.,* 1990), and iron-oxidising bacteria such as *Gallionella* and *Sphaerotilus* associated with biocorrosion processes (Emde *et al*., 1992; Sun *et al*., 2014).

2.3.2 Disinfectant type: chlorination and chloramination

In drinking water sytems that disinfect, disinfection is considered the most significant in terms of controlling microbial growth, maintaining water quality within DWDS and shaping the drinking water microbiome. Chlorine, chloramine and other chlorine compounds have long been successfully used for the control of microbial growth within DWDSs. Chlorine is a strong oxidising agent and it is known to be effective in injuring bacteria, thereby preventing and limiting their growth in the system (Vasconcelos *et al*., 1997). Chlorine exists in water as hypochlorite or hypochlorous acid which oxidises superficial biomolecules (membrane lipids and envelopes) as well as biomolecules within bacterial cells (enzymes and nucleic acids) resulting in cell death (Junli *et al*., 1997). For chlorine to be effective, it must be present at a sufficient concentrations and for a certain reaction time (Hwang *et al*., 2012b). Bertelli *et al*.

(2018) showed that high chlorine residuals reduced bacterial diversity causing a more homogenous bacterial population.

However, chlorine residuals (1mg Cl_2/L) are often insufficient to kill and remove attached biofilms, allowing a fraction of the bacterial population to escape disinfection (Norton and LeChevallier, 1997; Batté *et al*., 2003). Disinfectant residues are typically maintained at levels lower than the maximum guidelines (chlorine: 5.0 mg/L and chloramine: 3.0 mg/L) (WHO, 2011). Chlorine is present in most disinfected drinking water at concentrations of 0.2 – 1 mg/L and chloramine at 0.5 – 2 mg/L (WHO, 2011). The efficacy of disinfectants are influenced by biofilm formation and penetration of the disinfectant into biofilms as well as the presence of AOC and possible corrosion of the pipe surfaces (Norton and LeChevallier, 1997). It is important that a chlorine residual be maintained throughout the system to be effective. However, there are multiple factors affecting the depletion of chlorine in the distribution system such as, composition of pipe materials, biofilms, presence of organic matter in the bulk water, size of the system, hydraulic forces and residence times (Norton and LeChevallier, 1997; Chu *et al.,* 2005). In addition, disinfection can potentially form carcinogenic disinfection by-products (DBP) such as trihalomethanes (THM), haloacetic acids (HAA), halonitromethanes (HNM), haloacetonitriles (HAN), etc. (Wilczak *et al.,* 1996; Goslan *et al.,* 2009; Bougeard *et al.,* 2010).

Although chlorination is successful in initially reducing bacterial growth, many distribution systems, including South Africa, apply chloramination as a second disinfection strategy. Following an initial disinfection event, secondary disinfection aims to reduce microbial growth within the distribution system. In both cases, disinfection greatly alters the composition and structure of the microbial community (Prest *et al*., 2016a). Hwang *et al*. (2012b) observed that chlorination and chloramination treatments exerted strong selection pressures on the microbial community. Chloramines are typically applied in a 3:1 to 4:1 ratio of chlorine to ammonia. The ammonia associated with these ratios optimises the formation of monochloramines. An effective disinfectant should maintain its lethality by maintaining residual concentrations for longer distances downstream in the system. At the same concentrations of free chlorine, monochloramines have the same oxidising potential as free chlorine and are more stable and persistent in maintaining residual disinfectant throughout the DWDS (Norton and LeChevallier, 1997; Zhang *et al*., 2009). They decrease heterotrophic

bacteria and coliform growth as well as improve the taste and odour of drinking water (Neden *et al.,* 1992). Volk and LeChevallier (1999) showed that the consumption of chloramine was lower than that of chlorine. Chloramine is known to be more effective in reducing biofilm growth as it is not consumed by the biofilms extracellular polysaccharide matrix, as it is able to penetrate the biofilm and interact specifically with the DNA and certain amino acids such as tryptophan and those containing sulphur (LeChevallier *et al.,* 1990; Geldreich, 1996; Vikesland *et al.*, 2001). Chloramines can themselves produce THM and HAA and are prohibited in some European countries. However, these DBPs of chloramination are only produced in trace amounts and much lower levels than that observed for chlorine (Norton and LeChevallier, 1997; Goslan *et al.,* 2009; Bougeard *et al.,* 2010).

2.3.3   Water te**m**perature

Temperature is an important factor to consider when assessing water quality as it has been shown to have a significant effect on microbial growth kinetics and both chemical and microbial processes within drinking water systems (LeChevallier *et al.,* 1996; Zlatanović *et al.*, 2017). Temperatures can fluctuate dramatically within a single DWDS on a seasonal scale. Pinto *et al.*, (2014) observed temporal trends within a drinking water bacterial community, which corresponded to seasonal cycles. Seasonal trends are often observed in a single distribution system where an increase in water temperature, typically in the warmer months of spring and summer, results in an increase in bacterial abundance and richness (Pinto *et al.*, 2014). Water temperature can therefore affect the bacterial community composition as an increase in temperature may allow for a competitive advantage for specific bacterial groups. Temperatures from 15°C and above have been shown to increase the growth of nitrifying bacteria (Kirmeyer, 1995; Pintar and Slawson, 2003).

An increase in water temperature in the summer months is often associated with an increased possibility of bacterial growth problems. In addition, the total effects of temperature may be influenced by other factors such as presence of chlorine residual, BOM and shear forces (Ollos *et al.,* 2003; Ndiongue *et al.,* 2005). Ndiongue *et al.* (2005) demonstrated that temperature appeared to have little effect on the level of biofilm formation before the addition of chlorine. However, when BOM was added this had a significant impact on bacterial numbers. When BOM is absent, temperature appeared to have no affect and shear forces

seemed to play a more important role whereas when BOM was present, temperature was more important than shear (Ollos, 1998; Ollos *et al.,* 2003).

2.3.4 Pipeline material and distribution system infrastructure

It is well known that the type of pipe material has a direct effect on the water quality and resulting microbial communities (Niquette *et al.,* 2000). Pipe materials not only affect the biofilm formation but also the microbial community composition, diversity and richness (Yu *et al.,* 2010; Douterelo *et al.,* 2014). Wide varieties of pipe materials have been used worldwide, depending on the different distribution systems, the cost and availability of materials. A single distribution system may have a diverse range of pipe materials differing in composition and age. Typically, pipeline materials can be broken down into three groups: cementitious, metallic (copper and steel) and plastic (polyethylene and polyvinyl chloride) (Momba *et al.,* 2000; Yu *et al.,* 2010), all of which can be colonised by microorganisms. Certain pipe materials can modify and increase decay of disinfectant residuals leading to an increase in microbial regrowth through biofilm formation (Hallam *et al.,* 2001; Lehtola *et al.,* 2004; Yu *et al.,* 2010). Biofilms directly interact with the pipe surfaces, especially with cast iron pipes, and the resulting microbial growth can lead to microbially mediated corrosion (MIC).

2.3.5 Hydraulic forces

Among the many factors affecting the microbial community dynamics, many studies have investigated the effects of hydraulic forces on microbial growth (Percival *et al.,* 1999; Ollos *et al.,* 2003; Lehtola *et al.,* 2006). Hydraulic conditions play an integral role in the interactions between the 4 phases described by Liu *et al*. (2014) (bulk water, biofilms, loose deposits and suspended solids). The bacteria associated with biofilms and loose particles can become resuspended in the bulk water phase and carried through the system to the consumers tap. Periods of low water demand may result in lower flow velocities and stagnation causing an increase in residence time, particles to sediment and ultimately increases in microbial growth (Liu *et al.,* 2013a and 2013b). Conversely, when water consumption is high, flow velocity increases causing increased bacterial dispersion through sediment re-suspension and potential shearing of biofilms (Lehtola *et al.,* 2006; Douterelo *et al*., 2014; Prest *et al*., 2016a).

Increases in bacterial numbers in biofilms has been linked to increases in flow velocity or shear forces (Percival *et al.,* 1999; Ollos *et al.,* 2003; Lehtola *et al.,* 2006), whereas contradicting studies have demonstrated that an increase in flow velocity causes a reduction in the levels of biofilm formation (Donlan and Pipes, 1988). These contradicting results may be attributed to the vast differences in various distribution systems, differences in the biofilm age and disinfectant residual concentrations. Tsai (2006) showed that in conditions of no chlorine or lower chlorine concentrations, the growth rate of biofilm increased with increases in shear stress although the opposite relationship occurred at higher chlorine concentrations. In addition, fluctuations in flow velocities results in resuspension of sediments and detachment of biofilm, both contributing to increases in turbidity and release of metals and organic matter (Batte *et al*., 2003; Lehtola *et al.,* 2006; Liu *et al.,* 2013a and 2014). These alternating scenarios result in fluctuations in the abundance and composition of bacterial communities within the bulk water.

2.3.6   Residence time and water age

Distances from treatment to tap has a significant effect on the residence time. The residence time of water within some distribution systems can reach up to a few days, leading to stagnation and increased opportunities for microbial growth (Prest *et al*., 2013; Zlatanović *et al*., 2017). In addition, the variability in pipe diameters and fluctuations in flow velocity, caused by water consumption, influences residence time specifically within reservoirs (Prest *et al*., 2016a). Here, an increase in residence time/stagnation correlates to an increase in bacterial abundances thereby potentially increasing disinfectant decay (Lehtola *et al*., 2007; Ling *et al*., 2018; Nescerecka *et al*., 2018). Wang *et al*. (2014) observed that changes in the water chemistry associated with increased water age such as decreases in disinfectant residual and dissolvable oxygen and an increase in TOC caused significant shifts in the microbial community. In addition, although in premise plumbing, Ling *et al*. (2018) observed highly reproducible spatially structured communities following stagnation events.

2.4 Challenges experienced in controlling microbial growth

DWDS are complex aquatic environments and problems experienced cannot be dealt with in isolation. Many of the challenges experienced within DWDSs are microbially based (Liu *et al*., 2014), with some of the most common challenges associated with biofilm formation and microbial growth. The presence of biofilms and loose deposits act as sites for biomass

accumulation and microbial mediated corrosion (Liu *et* al., 2013a). Furthermore, these sites of high biomass may harbour potential pathogens, increase depletion of disinfectant residual and deteriorate taste and odours.

2.4.1 Presence of pathogens

The growth and persistence of pathogens is a primary concern in DWDSs. The presence of pathogens in drinking water systems are typically associated with biofilm formation as biofilms offer a favourable environment and multiple advantages to associated bacteria (Wingender and Flemming, 2004; Chao *et al*., 2015). They provide protection from disinfection and environmental stresses, nutrients and metabolic products are commonly shared and opportunities for horizontal gene transfer are improved within the drinking water community (Wingender and Flemming, 2011). As a result of these benefits, potential pathogens may reside in the biofilm where they are protected from residual disinfectant and are able to proliferate to higher abundances in the oligotrophic environment of the DWDS resulting in an increased public health risk (Chao *et al*., 2015).

Pathogens such as *Legionella pneumophila, Pseudomonas aeruginosa, Aeromonas hydrophila, Klebsiella pneumoniae* and *Mycobacteria* all have the ability to grow in the low nutrient environments of DWDS (Flemming *et al*., 2002; Wang *et al*., 2014; Prest *et al*., 2016a). In addition to bacterial pathogens, there are some protozoan pathogens such as *Cryptosporidium* and *Giardia*, which may survive treatment systems and grow in DWDSs (Szewzyk *et al*., 2000). Free-living amoebae are also of concern in drinking water systems as some members can cause infections of the eye and brain as well as support the intracellular growht of bacteria pathogens such as, *Legionella* and *Mycobacteria* (Thomas and Ashbolt, 2011; Wang *et al.,* 2014). Lastly, the presence of viruses such as norovirus, hepatitis and rotavirus in drinking water has been reported (Gall *et al*., 2015). Waterborne viruses have a range of different capsid protein structures and genome type and are able to persist in water for long periods of time (Gall *et al*., 2015).

2.4.2 Microbial influenced corrosion

Beech and Sunner, (2004) defined biocorrosion or microbial influenced corrosion (MIC) as the accelerated deterioration of metals due to the presence of microorganisms on their surfaces. Biocorrosion is a result of the interactions between bacterial cells and their

metabolites, abiotic corrosion products and the metal surface (Beech and Sunner, 2004). The development of biofilms on metal surfaces of drinking water pipelines can considerably alter the chemistry and kinetics of corrosion reactions on metal pipe surface leading to the potential acceleration or inhibition of corrosion (Sun *et al*., 2014). Problems associated with MIC are more significant in iron and steel pipelines (Camper *et al.,* 2014). As mentioned previously, distribution systems with iron pipelines showed significantly higher biofilm densities compared to other materials (Niquette *et al.,* 2000). In addition, Sun *et al*. (2014) observed that the abundance of corrosion associated bacteria were significantly higher in biofilms originating from surface water than ground water. Typically, the main types of bacteria involved in MIC in aquatic habitats include iron-oxidising/reducing bacteria, sulphate-reducing bacteria, sulphur-oxidising bacteria and manganese-oxidising bacteria (Emde *et al*., 1992; Beech and Sunner, 2004; Sun *et al*., 2014). Consequently, the resulting corrosion by-products may provide a nutrient source or interact with the disinfectant residuals, reducing disinfectant efficacy (Niquette *et al.,* 2000; Batté *et al*., 2003; Yu *et al.,* 2010).

Bacteria may accumulate in tubercles created by corrosion of the iron pipe material (LeChevallier *et al*., 1987). Here, bacterial cells may be protected from environmental stresses and disinfection. Emde *et al.* (1992) observed a higher variety of species in corrosion-induced-deposits and in the bulk water phase following extended periods of chlorination. They concluded that corrosion tubercles are able to sustain a diverse population of microorganisms including direct and opportunistic pathogens as well as creating a habitat for microorganisms that influence the water's taste and odour (e.g. *Actinomycetes* and fungi).

2.4.3   Decreased disinfectant residuals and disinfectant decay

The quality of drinking water ultimately changes as a response to a series of physical, chemical and biological processes occurring as water moves through the complex pipe system. As a direct response to the decrease in water quality, water suppliers aim to maintain residual concentrations of disinfectant to minimise the potential for microbial growth (Ramos *et al*., 2010). However, disinfection concentrations typically decrease with increasing distance in the DWDS, unless disinfection boosting is applied. The decrease and consumption of residual disinfectant depends on multiple factors including reactivity with water, pipe wall material, hydraulic flow, residence time, temperature, pH, organic matter, microorganisms and biofilms (Vasconcelos *et al*., 1997; Chandy and Angles, 2001; Ramos *et al*., 2010). The

decrease in disinfection residuals ultimately promotes microbial activity where residual chlorine may also indirectly support microbial growth by increasing AOC concentrations through organic carbon breakdown and the production of DBPs (Fish and Boxall, 2018). As a result of these factors, biofilm development and microbial growth cannot be avoided (Rossman *et al*., 1994; Prest *et al*., 2016a). Chlorine is a highly reactive chemical and readily reacts with a variety of inorganic and organic compounds thereby causing its gradual decrease in the distribution system (Vasconcelos *et al*., 1997).

Chloramine is also inherently unstable in water and can decay by itself even in the absence of organic matter. The decay of chloramine often results in the liberation of free-ammonia and thereby becomes an energy source for nitrifiers (Bal Krishna *et al.,* 2016 and 2018). Here, the formation of nitrite, nitrate, soluble microbial products and the release of extracellular polymeric substances during nitrification may lead to increased chloramine decay (Moradi *et al*., 2017). However, microbially induced decay of chloramine may not only be a result of ammonia oxidising bacteria through nitrification but potentially by other heterotrophic bacteria and soluble microbial products (Sawade *et al*., 2016; Bal Krishna *et al.,* 2012 and 2018). The decay of disinfectant throughout DWDS ultimately results in a decrease in water quality and is therefore a major concern for drinking water suppliers.

2.4.4   Nitrification in chloraminated DWDSs

The breakdown of disinfectants such as monochloramines, into their constituent compounds may also themselves act as a nutrient source and accelerate disinfectant decay. Chloramines are often added as secondary disinfectant when free chlorine residuals are difficult to maintain. Chloramines are produced from reactions between free chlorine and ammonia with monochloramine (NH_2Cl) being the most commonly used compound in drinking water treatment. However, chloramination potentially causes undesirable nitrification, resulting in operational problems for many drinking water utilities (Regan *et al*., 2002). Bacterial nitrification can then lead to the rapid depletion of monochloramine residual concentrations (Cunliffe, 1991; Berry *et al*., 2006).

The introduction of ammonia into the system provides a potential source of nitrogen either by excess ammonia or through chloramine decay (Zhang *et al*., 2009; Bal Krishna *et al*., 2016). This promotes the growth of nitrifying microorganisms represented by different species of

both bacteria and archaea (Belser, 1979; Nicol and Schleper, 2006). Bacterial nitrification in the DWDS causes an increase in nitrite and nitrate levels impacting both water quality and infrastructure through MIC (Zhang *et al*., 2009; Wang *et al.,* 2014). In addition, nitrification results in the consumption of dissolvable oxygen and a decrease in pH (Kirmeyer *et al.,* 1995; Zhang *et al*., 2009). Increased nitrite in the system also rapidly decreases the free chlorine and is also further oxidised by chloramine leading to an accelerated decrease in residual chloramine (Wolfe *et al.,* 1990; Cunliffe, 1991). The loss of chloramine residual leads to an increase in heterotrophic bacterial concentrations and biofilm accumulation resulting in an increased potential for re-growth events within the distribution system (Kirmeyer *et al.,* 1995; Norton and LeChevallier, 1997; Pintar and Slawson, 2003).

Microbial nitrification is a two-step process: firstly ammonia (NH_4^+) is oxidised to nitrite (NO_2^-) by ammonia oxidising bacteria and archaea (AOB and AOA respectively) (van der Wielen *et al*., 2009). Secondly, nitrite is further oxidised to nitrate (NO_3^-) by nitrite oxidising bacteria (NOB) (Wolfe *et al.,* 1990; Cunliffe, 1991; Francis *et al.*, 2005). The free ammonia, nitrites and nitrates then serve as an energy source for ammonia-oxidising bacteria (AOB) and nitrite oxidising bacteria (NOB) (Kirmeyer *et al.,* 1995; Pintar and Slawson, 2003).

Nitrifying bacteria are chemolithothrophic bacteria. Common bacterial genera involved in ammonia oxidation typically include members of *Betaproteobacteria,* genera *Nitrosomonas and Nitrosospira* (Kowalchuk and Stephen, 2001; Zhang *et al*., 2009). In addition, multiple studies have identified ammonia oxidising archaea (AOA) including members from the phylum *Thaumarchaeota* (You *et al*., 2009; Spang *et al*., 2010; Stahl and de la Torre, 2012). Nitrite oxidising bacteria associated with aquatic environments include members from the genus *Nitrobacter* belonging to the *Alphaproteobacteria* as well as members from the genus *Nitrospira* (Zhang *et al*., 2009). However, recent studies have revealed the presence of the complete ammonia-oxidising (comammox) *Nitrospira*-like bacteria in DWDSs, which is capable of the completing the full oxidation of ammonia to nitrate (Diams *et al*., 2015; van Kessel *et al*., 2015; Pinto *et al*., 2016; Wang *et al*., 2017). In light of this discovery, it is possible that comammox bacteria play a more significant role in nitrification in drinking water systems than previously thought (Bautista-de los Santos *et al*., 2016; Tarari *et al*., 2017).

3. Exploring microbial ecology in drinking water distribution systems

3.1 Direct measurements and enumeration of bacterial concentrations

Techniques used to study microbial ecology can typically be divided into culture-dependant and culture-independent. Typically, culture-dependent techniques involve isolating microbes from the environment and conducting biochemical tests (Clark *et al*., 2018). In drinking water systems, traditionally heterotrophic plate counts (HPC) have been used to determine bacterial concentrations. This culture-based test relies on the fact that heterotrophic bacteria, yeast and moulds require organic carbon for growth. However, only a small portion of metabolically active microorganisms in the water sample may grow. Although this technique still provides valuable information regarding growth rates, metabolic requirements and optimal growth conditions of heterotrophic bacteria, as a stand-alone method its application in microbial ecology is limited as it leads to a great underestimation of the total microbial community present within the DWDS (Liu *et al.,* 2013a; Chaio *et al.,* 2014; Clark *et al*., 2018). HPC test are still employed today as indicators for the effectiveness of the water treatment process and therefore are indirect indicators of pathogen removal and water safety.

To overcome the disadvantages of HPC in terms of bacterial enumeration, total cell counts (TCC) can be used. This method employs membrane filtration, fluorescent dye staining and microscopic counting to determine bacterial cell numbers (Boe-Hansen *et al.,* 2002). Fluorescent staining involves DAPI (4',6-diamidino-2-phenylindole) staining, which binds to double-stranded DNA and can pass through intact cell membranes giving an indication of the proportion of live cells in the samples. Stained cells can then be visualised microscopically.

More recently, flow cytometry (FCM) has shown great potential for monitoring total cell counts in drinking water systems (Berney *et al.,* 2008; Hammes *et al*., 2008; Prest *et al*., 2016b; Nescerecka *et al*., 2018). FCM is simple and rapid as well as more sensitive and accurate. Fluorescent labelling of nucleic acids allows for direct enumeration of total cell concentrations as well as detection of specific cellular features linked to cell viability (Hammes *et al*., 2008). Cells can be differentiated as intact or damaged on the basis of their cell membrane integrity. Live cells have intact membranes that exclude a variety of dyes. Therefore, dyes used for the visualisation of total intact cells include SYBR Green, SYTO9 and thiozole orange (TO) as they are able to penetrate intact cell membranes and bind with

double stranded DNA. In addition, for the visualisation of damaged cells, stains such a propidium iodide (PI) are used. PI is a membrane impermeant dye that is typically excluded from viable cells and therefore can penetrate damaged or compromised cell membranes.

Alternatively, the level of biologically active bacteria within the sample can be measured using adenosine triphosphate (ATP) as high ATP numbers correlate to high bacterial numbers (Hammes *et al*., 2010; Liu *et al.,* 2013a). In addition, concentrations of AOC within the sample are used to determine the potential for microbial growth (Momba *et al.,* 2000; Ndiongue *et al.,* 2005). Hammes *et al*. (2008) compared data obtained from FCM with conventional HPC as well as ATP concentrations from different drinking water treatment processes. They observed that FCM showed clear advantages over HPC and ATP as it detects cells irrespective of culturability and that ATP measurements are often affected by extracellular ATP. In conclusion they suggest that total cell enumeration through FCM is a valuable tool in monitoring water quality treatment and distribution.

3.2 Bacterial identification and community analysis

Of the methods discussed, none provides an indication of the microbial species present, their abundance or contribution to the total community within the distribution system. The abovementioned methods give an indication of cell numbers and cell viability. As mentioned, HPC is used for the detection of heterotrophic bacteria in DWDS, the resulting numbers are a great underestimation of the microbial community present as only culturable microorganisms are identified. For this reason, culture-independent and molecular methods have been developed.

Culture-independent techniques bypass the need to culture and isolate microbes from the collected environmental samples. The majority of these techniques begin with DNA extraction directly from the sample where the genetic material may include genes, genomes, transcripts and transcriptomes from potentially all organisms present in the samples (Clark *et al*., 2018). The next step may include amplification of nucleic acids via the polymerase chain reaction (PCR) where the choice of amplicon target depends on the research question. Typically, genetic markers such as the 16S rRNA gene is targeted as it provides information on the evolutionary identity of the organism or alternatively, functional markers may be

targeted to understand the functional capabilities of an organism (e.g. *amoA* genes in ammonia oxidisers) (Clark *et al*., 2018).

16S rRNA gene amplification and analysis includes techniques such as denaturing gradient gel electrophoresis (DGGE), terminal restriction fragment length (T-RFLP), cloning and 16S rRNA gene sequencing (Hoefel *et al.,* 2005). These techniques make use of differences in nucleotide composition and molecular weight to create a "fingerprint" that can be compared between samples. T-RFLP has been used to examine the DWDS microbial community in several studies (Hwang *et al.,* 2012a; Hwang *et al.,* 2012b; Wang *et al.,* 2012; Zhang *et al.,* 2012). Another fingerprinting method used is 16S rRNA single-strand conformation polymorphism (SSCP) based on extracted DNA and RNA, which reveals bacterial species with relative high abundances (Schmeisser *et al.,* 2003; Eichler *et al.,* 2006; Henne *at al.,* 2012). These molecular fingerprints allow for evaluation of relative abundance of a species, species richness and community composition (Henne *at al.,* 2012) but they are unable to represent a complete representation of microbial diversity and community composition due to limited throughput (Sun *et al.,* 2014).

3.2.1   Understanding microbial ecology through next generation sequencing

3.2.1.1 16S rRNA gene profiling

As it is not possible to eliminate bacteria from drinking water, it is crucial to understand the extent of the microbial community within the system, that is, which species are present, their relative abundances and how treatment and the distribution system parameters shape the microbial community structure (Pinto *et al.,* 2012a). The development of high-throughput and deep DNA sequencing such as 454 pyrosequencing, Ion Torrent and Illumina platforms have greatly improved our understanding of the drinking water microbiome and allow for the sequencing of millions of amplicons in parallel (Bautista-de los Santos *et al*., 2016). Here, 16S rRNA gene amplicon sequencing can be targeted specifically for bacteria with no required available reference sequences and it can be employed in cases where DNA concentrations are low or of poor quality (Salipante *et al*., 2014). Next generation sequencing (NGS) targeting the 16S rRNA gene has been used in several recent studies. These studies highlighted the impact of seasonal changes (Pinto *et al*., 2014; Prest *et al*., 2016b), treatment and disinfectant type (Gomez-Alvares *et al*., 2012; Hwang *et al*., 2012b, Bautista-de los Santos *et al.,* 2016; Bertelli *et al*., 2018; Bruno *et al*., 2018), process operations (Pinto *et al*.,

2012a; Lautenschlager *et al*., 2014; Wu *et al*., 2015), hydraulic conditions and flushing (Douterelo *et al*., 2014; El-Chakhtoura *et al.,* 2018), pipe material and system infrastructure (Yu *et al*., 2012; Wang *et al*., 2014) and water age and residence time (Wang *et al*., 2014; Prest *et al.*, 2016a; Zlatanovic *et al*., 2017) on the bacterial community composition and structure.

Multiple studies have made use of 454 pyrosequencing for microbial community and diversity analyses of both bulk water and biofilms in DWDS (Hong *et al.,* 2010; Kwon *et al.,* 2011; Pinto *et al.,* 2012a,b; Zhang *et al.,* 2012; Chaio *et al.,* 2014; Douterelo *et al.,* 2014; Sun *et al.,* 2014). All these studies demonstrate the ability of pyrosequencing to characterise microbial communities within DWDS revealing how the disinfectant regimes and various environmental factors influence and change the community composition and diversity. Here, taxonomic identification is improved and rare species within the community are identified.

More recently, the development of Illumina MiSeq technology enabled high-resolution characterisation of microbial communities (Caporaso *et al*., 2012). Fadrosh *et al.* (2014) described a dual-indexing amplification and sequencing approach to assess the composition of the microbial community using the Illumina MiSeq platform. Illumina is now replacing 454 pyrosequencing as the method of choice for microbial community related studies as this approach provides a more cost effective and flexible sequencing option (Kozich *et al*., 2013; Salipante *et al*., 2014). Here, community profiling studies using Illumina MiSeq, have revealed shifts in the microbial communities due to sampling location, treatment processes, distribution as well as temporal variations (Baron *et al*., 2014; Roeselers *et al*., 2015; Wu *et al*, 2015).

Community profile data generated from this high throughput sequencing provides insight into the diversity within the drinking water environment. Typically, microbial diversity is not evenly distributed as different sites contain different microbial communities at varying abundances. Therefore, measuring the differences within and between communities can improve the understanding of how microbial diversity is distributed within drinking water systems. Researchers are able to observe the species present and their abundance in the community within specific site (alpha diversity) as well as compare the species diversity

between different sites or the degree of community differentiation (beta diversity) (Tuomisto, 2010).

3.2.1.2 Metagenomic approach

There has been a significant increase in research regarding the microbial ecology of drinking water systems. However, many studies have focused solely on the taxonomic composition of the microbial community and very little in their functional and metabolic potential (Douterelo *et al*., 2018b). Metagenomics provides access to the functional gene composition within a microbial community, thereby providing a more comprehensive description of the genomes contained within a specific environment than phylogenetic surveys such as those based 16S rRNA gene amplification (Thomas *et al*., 2012). Shotgun metagenomic sequencing is a relatively new and robust approach to environmental sequencing, shedding light on microbial community biodiversity and function. This approach has uncovered the extensive biodiversity within microbial communities as well as microbial ecology in terms of the functions behind the interactions between individuals and their environment (Sharpton, 2014).

Here DNA is extracted directly from the environment, from all cells in the community. Total genomic DNA is then sheared into smaller fragments that are then sequenced individually. The resulting reads consist of DNA sequences from various regions of the genomes of different individuals within the community. Some reads will be taxonomically informative where as others will provide insight into the biological functions encoded in the genomes of the community members (Sharpton, 2014). Metagenomic reads containing protein coding sequences are identified and the potential gene function is predicted by comparing sequences to databases of genes, proteins or protein families and metabolic pathways. This potentially produces a functional profile of the community, revealing functions that are associated with a specific environment (Sharpton, 2014).

Using this approach multiple studies have revealed differences in community function between biofilms and bulk water (Douterelo *et al*., 2018b), between different sand filters in drinking water treatment (Oh *et al*, 2018), between different temperatures and stagnation settings (Dai *et al*., 2018) and between drinking water biofilms formed on stainless steel and plastics (Chao *et al*., 2015). Furthermore, using a metagenomic approach Chao *et al*. (2013) were able to show differences in metabolic functions between the microbial communities in

source water versus treated water. They identified an increase genes involved in protective functions such as glutathione synthesis genes linked to oxidative stress and detoxification in treated water, suggesting that bacterial surviving disinfection may have higher chlorine resistance.

Furthermore, the construction of metagenome assembled genomes (MAGs) from metagenomic data can be applied to investigate individual genomes of unculturable microorganisms from various ecosystems and reconstruct partially-complete genomes of dominant members of the community (Allen and Banfield, 2005; Chao *et al.*, 2015). This process of binning involves the sorting of DNA sequences into groups or bins, representing an individual microorganism or closely related microorganisms (Thomas *et al.*, 2012). Using Illumina metagenomic data, Chao *et al.*, (2015) were able to identify some of the key metabolic functions of the drinking water biofilms as well as reconstruct partial genome of *Bradyrhizobiaceae*-like bacterium. The reconstruction of MAGs has also lead to the discovery of comammox *Nitrospira*-like bacteria from multiple drinking water systems (Diams *et al.*, 2015; Pinto *et al.*, 2016; Wang *et al.*, 2017).

4. Conclusion

Drinking water utilities aim to produce water that is microbially safe, through treatment and distribution and finally to the consumers tap. Based on current information, it is evident that consecutive treatment operations employed by DWTPs not only improve water quality, but have also have a significantly impact on the drinking water microbiome. However, as drinking water moves through the DWDS, water quality is likely to deteriorate as a result of interactions with chemical, physical and biological factors inherent to each DWDS. Therefore it may be beneficial to drinking water utilities to understand the impact of these multiple factors on the microbial ecology within drinking water systems.

It is apparent that of all the factors influencing microbial community dynamics, no one factor acts in isolation. The extent to which one factor influences the microbial community, is itself influenced (positively or negatively) by another and together they will have a certain effect on the microbial dynamics within a drinking water system. Drinking water systems (i.e. treatment and distribution) are multi-dimensional complex systems where all contributing factors will have an overall combined effect on the microbial community and ultimately

water quality. The complexity of interactions between the contributing factors, differing source waters, the variability between distribution systems and the difficulty in standardisation of those systems may explain the various conflicting results observed in literature. Few studies have explored the interaction of different physico-chemical factors in the drinking water environment and their impact, individually or in combination, on the microbial community composition and structure.

Microbial ecology aims to determine what factors shape and maintain the composition and structure of microbial communities by elucidating the origin of the community, understanding the community's stability and resilience, identify the factors that stabilize or restrict the community and lastly to determine the diversity and link it to community function. Studies linking microbial diversity (who is there?) and function (what are they doing?) are limited. Here, the use of metagenomics and transcriptomics can provide a detailed picture of the functional and metabolic activity of a community at a specific time point, and can be used to examine changes to this in response to environmental disturbances or existing conditions.

In addition, studies investigating drinking water systems need to be standardised, adapted and optimized for application in different systems. Microbial communities need to be observed, sampled and described across appropriate spatial and temporal scales with sufficient replications, while monitoring environmental and physico-chemical parameters. Here, comprehensive and representative investigations will contribute to the fast growing body of information explaining the microbial ecology of drinking water systems and may contribute to the improvement of water management operations in the future. Ultimately, complete understanding of the microbial community will aid in setting up a potential predictive framework, which will help in eliminating microbial risks as well as upgrade water quality monitoring methods, making them more resource efficient and sustainable.

5. References

1. Albers, C. N., Ellegaard-Jensen, L., Harder, C. B., Rosendahl, S., Knudsen, B. E., Ekelund, F. and Aamand, J. (2015). Groundwater chemistry determines the prokaryotic community structure of waterworks sand filters. *Environmental Science and Technology*. 49(2): 839-846.

2. Allen, E. E. and Banfield, J. F. (2005). Community genomics in microbial ecology and evolution. *Nature Reviews*. 3: 489-498.

3. Bal Krishna, K., Sathasivan, A. and Sarker, D. C. (2012). Evidence of soluble microbial products accelerating chloramine decay in nitrifying bulk water samples. *Water Research*. 46(13): 3977-3988.

4. Bal Krishna, K. C., Bhullar, G. S., Sathasivan, A. and Henderson, R. (2016). Effectiveness of re-chloramination to control nitrification in chloraminated bulk waters. *Desalination and Water Treatment*. 57(34): 15970-15978.

5. Bal Krishna, K., Ginige, M. P. and Sathasivan, A. (2018). Is nitrite from nitrification the only cause of microbiologically induced chloramine decay? *Microbiology Australia*. 39(3): 145-148.

6. Baron, J. L., Vikram, A., Duda, S., Stout, J. E. and Bibby, K. (2014). Shift in the microbial ecology of a hospital hot water system following the introduction of an on-site monochloramine disinfection system. *PLoS One*. 9(7): e102679.

7. Batté, M., Appenzeller, B. M. R., Grandjean, D., Fass, S., Gauthier, V., Jorand, F., Mathieu, L., Boualam, M., Saby, S. and Block, J. C. (2003). Biofilms in drinking water distribution systems. *Reviews in Environmental Science and Biotechnology*. 2(2-4): 147-168.

8. 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.

9. Beech, I. B. and Sunner, J. (2004). Biocorrosion: towards understanding interactions between biofilms and metals. *Current Opinion in Biotechnology*. 15: 181-186.

10. Belser, L. W. (1976). Population Ecology of Nitrifying Bacteria. *Annual Review of Microbiology*. 33: 309 -333.

11. Berney, M., Vital, M., Hulshoff, I., Weilenmann, H. U., Egli, T. and Hammes, F. (2008). Rapid, cultivation-independent assessment of microbial viability in drinking water. *Water. Research.* 42(14): 4010-4018.

12. Berry, D., Xi, C. and Raskin, L. (2006). Microbial ecology of drinking water distribution systems. *Current Opinion Biotechnology.* 17: 297-302.

13. Bertelli, C., Courtois, S., Rosikiewicz, M., Piriou, P., Aeby, S., Robert, S., Loret, J. F. and Greub, G. (2018). Reduced chlorine in drinking water distribution systems impacts bacterial biodiversity in biofilms. *Frontiers in Microbiology.* 9: 2520.

14. 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.

15. Bondank, E. N., Chester, M. V. and Ruddell, B. L. (2018). Water Distribution System Failure Risks with Increasing Temperatures. *Environmental Science and Technology.* 52(17): 9605-9614.

16. Bougeard, C. M. M., Goslan, E. H., Jefferson, B and Parson, S. A. (2010). Comparison of the disinfection by-product formation potential of treated waters exposed to chlorine and monochloramine. *Water Research.* 44:729-740.

17. 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 a urbanized area, Milan (Italy). *Frontiers in Microbiology.* 9: 2557.

18. Camper, A. K. (2014). Organic matter, pipe material, disinfectants and biofilms in distribution systems. In *Microbial Growth in Drinking Water Supplies: Problems, Causes, Control and Research needs.* Van der Kooij, D and van der Wielen, P. W. J. J. (eds). London, UK: IWA Publishing, pp. 77-94.

19. Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M. and Gormley, N. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal.* 6: 1621–1624.

20. Chandy, J. P. and Angles, M. L. (2001). Determination of nutrients limiting biofilm formation and the subsequent impact on disinfectant decay. *Water Research.* 35(11): 2677-2682.

21. Chao, Y., Ma, L., Yang, Y., Ju, F., Zhang, X. X., Wu, W. M. and Zhang, T. (2013). Metagenomic analysis reveals significant changes of microbial compositions and protective functions during drinking water treatment. *Scientific Reports*. 3: 3550.

22. Chao, Y., Mao, Y., Wang, Z. and Zhang, T. (2015). Diversity and functions of bacterial community in drinking water biofilms revealed by high-throughput sequencing. *Scientific Reports*. 5: 10044.

23. 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.

24. Chu. C., Lu. C. and Lee. C. (2005). Effects of inorganic nutrients on the regrowth of heterotrophic bacteria in drinking water distribution systems. *Journal Environmental Management*. 74: 255-263.

25. Clark, D. R., Ferguson, R. M., Harris, D. N., Matthews Nicholass, K. J., Prentice, H. J., Randall, K. C., Randell, L., Warren, S. L. and Dumbrell, A. J. (2018). Streams of data from drops of water: 21st century molecular microbial ecology. *Wiley Interdisciplinary Reviews: Water*. 5(4): e1280.

26. Cunliffe, D. A. (1991). Bacterial nitrification in chloraminated water supplies. *Applied and Environmental Microbiology*. 57(11): 3399-3402.

27. Dai, D., Rhoads, W. J., Edwards, M. A. and Pruden, A. (2018). Shotgun metagenomics reveals taxonomic and functional shifts in hot water microbiome due to temperature setting and stagnation. *Frontiers in Microbiology*. 9: 2695.

28. Daims, H., Lebedeva, E. V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A. and Kirkegaard, R. H. (2015). Complete nitrification by Nitrospira bacteria. *Nature*. 528(7583): 504.

29. Donlan, R. M. and Pipes, W. O. (1988). Selected drinking water characteristics and attached microbial population density. *Journal of American Water Works Association*. 80(11): 70-76.

30. Douterelo, I., Sharpe, R. L. and Boxall, J. B. (2013). Influence of hydraulic regimes on bacterial community structure and composition in an experimental drinking water distribution system. *Water Research*. 47(2): 503-516.

31. Douterelo, I., Husband, S. and Boxall, J. B. (2014). The bacterial composition of biomass recovered by flushing and operational drinking water distribution system. *Water Research*. 54: 100-114.

32. Douterelo, I., Jackson, M., Solomon, C. and Boxall, J. (2017). Spatial and temporal analogies in microbial communities in natural drinking water biofilms. *Science of the Total Environment*. 581: 277-288.

33. Douterelo, I., Sharpe, R. L., Husband, S., Fish, K. E. and Boxall, J. B. (2018a). Understanding microbial ecology to improve management of drinking water distribution systems. *Wiley Interdisciplinary Reviews: Water*. p.e01325.

34. Douterelo, I., Calero-Preciado, C., Soria-Carrasco, V. and Boxall, J. B. (2018b). Whole metagenome sequencing of chlorinated drinking water distribution systems. *Environmental Science: Water Research and Technology*. 4(12): 2080-2091.

35. Edzwald, J. K. (1993). Coagulation in drinking water treatment: particles, organics and coagulants. *Water Research and Technology*. 27(11): 21-35.

36. Eichler, S., Christen, R., Holtje, C., Westphal, P., Botel, J., Brettar, I., Mehling, A. and Hofle, 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.

37. 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: 180e190.

38. El-Chakhtoura, J., Saikaly, P. E., van Loosdrecht, M. C. and Vrouwenvelder, J. S. (2018). Impact of Distribution and Network Flushing on the Drinking Water Microbiome. *Frontiers in Microbiology*. 9.

39. Emde, K. M. E., Smith, D. W. and Facey, R. (1992). Initial investigation of microbially influenced corrosion (MIC) in a low temperature water distribution system. *Water Research*. 26(2): 169-175.

40. Fadrosh, D. W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R. M. and Ravel, J. (2014). An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*. 2(6): 1-7.

41. Farkas, A., Dragan-Bularda, M., Muntean, V., Ciataras, D. and Tigan, S. (2013). Microbial activity in drinking water-associated biofilms. *Central European Journal of Biology*. 8(2): 201-214.

42. Fish, K. and Boxall, J. B. (2018). Biofilm microbiome (re) growth dynamics in drinking water distribution systems are impacted by chlorine concentration. *Frontiers in Microbiology*. 9: 2519.

43. Flemming, H. C., Percival, S. I. and Walker, J. T. (2002). Contamination potential of biofilms in water distribution systems. *Water Supply.* 2: 271-280.

44. Flemming, H. C., Bendinger, B., Exner, M., Gebel, J., Kistemann, T., Schaule, G., Szewzyk, U., Wingender, J., van der Kooij, D. and van der Wielen, P. W. J. J. (2014). The last meters before the tap: where drinking water quality is at risk (pp. 207-232). IWA Publishing: London.

45. Francis, C. A., Roberts, K. J., Beman, J. M., Santoro, A. E. and Oakley, B. B. (2005). Ubiquity and diversity of ammonia-oxidising archaea in water columns and sediments of the ocean. *Proceedings of the National Academy of Sciences.* 102(41): 14683-14688.

46. Gall, A. M., Mariñas, B. J., Lu, Y. and Shisler, J. L. (2015). Waterborne viruses: A barrier to safe drinking water. *PLoS Pathogens.* 11(6): e1004867.

47. Geldreich E. E. (1996). Microbial quality of water supply in distribution systems. CRC Press, Boca Raton, Fla.

48. Gomez-Alvarez, V., Revetta, R. P. and Santo Domingo, J. W. (2012). Metagenomic analysis of drinking water receiving different disinfection treatments. *Applied and Environmental Microbiology.* 78(17): 6095-6102.

49. Gomez-Alvarez, V., Humrighouse, B. W., Revetta, R. P. and Santo Domingo, J. W. (2015). Bacterial composition in a metropolitan drinking water distribution system utilizing different source waters. *Journal of Water and Health.* 13(1): 140-151.

50. Goslan, E. H., Krasner, S. W., Bower, M., Rocks, S. A., Holmes, P., Levy, L. S. and Parsons, S. A. (2009). A comparison of disinfection by-products found in chlorinated and chloraminated drinking waters in Scotland. *Water Research.* 43: 4698-4706.

51. Gülay, A., Musovic, S., Albrechtsen, H. J., Al-Soud, W. A., Sorensen, 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.* doi:10.1038/ismej.2016.16

52. Hallam, N. B., West, J. R., Forester, C. F. and Simms, J. (2001). The potential for biofilm growth in water distribution systems. *Water Research.* 35(17): 4063-4071.

53. 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.

54. Hammes, F., Goldschmidt, F., Vital, M., Wang, Y. and Egli, T. (2010). Measurement and interpretation of microbial adenosine tri-phosphate (ATP) in aquatic environments. *Water Research*. 44(13): 3915-3923.

55. Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C. and Martiny, J. B. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology*. 10(7): 497.

56. Henne, K., Kahisch, L., Brettar, I. and Holfe, M. G. (2012). Analysis of structure and composition of bacterial core communities in mature drinking water biofilms and bulk water of a citywide network in Germany. *Applied and Environmental Microbiology*. 78(10): 3530-3538.

57. Hoefel, D., Monis, P. T., Grooby, W. L., Andrews, S. and Saint, C. P. (2005). Culture-independent techniques for rapid detection of bacteria associated with the loss of chloramine residual in a drinking water system. *Applied and Environmental Microbiology*. 71(11): 6479-6488.

58. Hong, P., Hwang, C., Ling, F., Anderson, G. L., LeChevallier, M. W. and Liu, W. (2010). Pyrosequencing analysis of bacterial biofilm communities in water meters of a drinking water distribution system. *Applied and Environmental Microbiology*. 76(16): 5631-5635.

59. Hunter, P. R, MacDonald, A. M. and Carter, R. C. (2010). Water supply and health. *PLOS. Med.* 7(11): e1000461.

60. Hwang, C., Ling, F., Andersen, G. L., LeChevallier, M. W. and Liu, W. (2012a). Evaluation of methods for the extraction of DNA from drinking water distribution system biofilms. *Microbes and Environments*. 27(1): 9-18.

61. Hwang, C., Ling, F., Andersen, G. L., LeChevallier, M. W. and Liu, W. (2012b). 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.

62. Junli, H., Li, W., Nanqi, R. and Fang, M. (1997). Disinfection effect of chlorine dioxide on bacteria in water. *Water Research*. 31(3): 607-613.

63. 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.

64. Kowalchuk, G. A. and Stephen, J. R. (2001). Ammonia-oxidising bacteria: A model for molecular microbial ecology. *Annual Review of Microbiology*. 55: 485-529

65. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing

amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology.* AEM-01043.

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

67. Lautenschlager, K., Hwang, C., Ling, F., Liu, 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.

68. LeChevallier M. W., Babcock T. W. and Lee R. G. (1987). Examination and characterisation of distribution system biofilms. *Applied and Environmental Microbiology.* 53: 2714–2724.

69. LeChevalier, M. W., Cawthon, C. D. and Lee, R. G. (1988). Inactivation of biofilm bacteria. *Applied and Environmental Microbiology.* 54: 2492-2499.

70. LeChevallier, M. W., Lowry, C. D. and Lee, R. G. (1990). Disinfecting biofilms in a model distribution system. *Journal of American Water Works Association.* 82(7): 87-99.

71. LeChevallier, M. W., Shaw, N. J. and Smith, D. B. (1996). Factors limiting microbial growth in distribution systems: full-scale experiments. *Journal of American Water Works Association.*

72. Lehtola, M. J., Miettinen, I. T., Vartiainen, T. and Martikainen, P. J. (2002). Changes in content of microbially available phosphorus, assimilable organic carbon and microbial growth potential during drinking water treatment processes. *Water Research.* 36: 3681-3690.

73. Lehtola, M. J., Miettinen, I. T., Keinanen, M. M., Kekki, T. K., Laine, O., Hirvonen, A., Vartiainen, T. and Martikainen, P. J. (2004). Microbiology, chemistry and biofilm development in a pilot drinking water distribution system with copper and plastic pipes. *Water Research.* 38: 3769-3779.

74. Lehtola, M. J., Laxander, M., Miettinen, I. T., Hirvonen, A., Vartiainen, T. and Martikainen, P. J. (2006). The effects of changing water flow velocity on the formation of biofilms and water quality in pilot distribution system consisting of copper or polyethylene. *Water Research.* 40: 2151-2160.

75. Lehtola, M. J., Miettinen, I. T., Hirvonen, A., Vartiainen, T. and Martikainen, P.J. (2007). Estimates of microbial quality and concentration of copper in distributed drinking water are

highly dependent on sampling strategy. *International Journal of Hygiene and Environmental Health*. 210(6): 725-732.

76. Li, C., Ling, F., Zhang, M., Liu, W. T., Li, Y. and Liu, W. (2017). Characterization of bacterial community dynamics in a full-scale drinking water treatment plant. *Journal of Environmental Sciences*. 51: 21-30.

77. 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.

78. 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.

79. 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.

80. Little, A. E., Robinson, C. J., Peterson, S. B., Raffa, K. F. and Handelsman, J. (2008). Rules of engagement: interspecies interactions that regulate microbial communities. *Annual Review of Microbiology*, 62: 375-401.

81. Liu, G., Verbeck, J. Q. J. C. and Van Dijk, J. C. (2013a). Bacteriology of drinking water distribution systems: an integral and multidimensional review. *Applied and Environmental Microbiology*. 97: 9265-9276.

82. Liu, G., Lut, M. C., Verberk, J. Q. J. C. and Van Dijk, J. C. (2013b). A comparison of additional treatment processes to limit particle accumulation and microbial growth during drinking water distribution. *Water Research*. 47(8): 2719-2728.

83. Liu, G., Bakker, G. L., Vreeberg, J. H. G., Verberk, J. Q. J. C., Medama, G. J., Liu, M. C. and Van Dijk, J. C. (2014). Pyrosequencing reveals bacterial communities in unchlorinated drinking water distribution system: An integral study of bulk water, suspended solids, loose deposits and pipe wall biofilm. *Environmental Science and Technology*. 48: 5467-5476.

84. Liu, J., Zhao, R., Zhang, J., Zhang, G., Yu, K., Li, X. and Li, B. (2018). Occurrence and fate of ultramicrobacteria in a full-scale drinking water treatment plant. *Frontiers in Microbiology*. 9.

85. Matilainen, A., Vepsäläinen, M. and Sillanpää. (2010). Natural organic matter removal by coagulation during drinking water treatment: A review. *Advances in Colloid and Interface Science*. 159: 189-197.

86. Miettinen, I. T., Vartiainen, T. and Martikainen, P. J. (1997). Phosphorus and bacterial growth in drinking water. *Applied and Environmental Microbiology*. 63(8): 3242-3245.

87. Momba, M. N. B., Kfir, R., Venter, S. N. and Cloete, T. E. (2000). An overview of biofilm formation in the distribution systems and its impact on the deterioration of water quality. *Water SA*. 26(1): 59-66.

88. Moradi, S., Liu, S., Chow, C. W., van Leeuwen, J., Cook, D., Drikas, M. and Amal, R. (2017). Developing a chloramine decay index to understand nitrification: A case study of two chloraminated drinking water distribution systems. *Journal of Environmental Sciences*. 57: 170-179.

89. Ndiongue, S., Huck, P. M. and Slawson, R. M. (2005). Effects of temperature and biodegradable organic matter on control of biofilms by free chlorine in a model drinking water distribution system. *Water Research*. 39: 953-964.

90. Neden, D. G., Jones, R. J., Smith, J. R., Kirmayer, G. J. and Foust, G. W. (1992). Comparing chlorination and chloramination for controlling bacterial regrowth. *Journal of American Water Works Association*. 80-88.

91. Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P. and Ferrenberg, S. (2013). Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews*. 77(3): 342-356.

92. 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.

93. Nicol, G. W. and Schleper, C. (2006). Ammonia-oxidising Crenachaeota: important players in the nitrogen cycle? *Trends in Microbiology*. 12(5): 207-212.

94. Niquette, P., Servais, P. and Savoir, R. (2000). Impacts of pipe materials on densities of fixed bacterial biomass in drinking water distribution system. *Water Research*. 34(6): 1952-1956.

95. Norton, C. D and LeChevallier, M. W. (1997). Chloramination: its effect on distribution system water quality. *Journal of American Water Works Association*. 89(7): 66-77.

96. Norton, C. D and LeChevallier, M. W. (2000). A pilot study of bacteriological population changes through potable water treatment and distribution. *Applied and Environmental Microbiology*. 66(1): 268-276.

97. Oh, S., Hammes, F. and Liu, W. T. (2018). Metagenomic characterization of biofilter microbial communities in a full-scale drinking water treatment plant. *Water Research*. 128: 278-285.

98. Ollos, P. J. (1998). Effects of drinking water biodegradability and disinfectant residual on bacterial growth. Doctoral Thesis, University of Waterloo, Ontario.

99. Ollos, P. J, Huck, P. M. and Slawson, R. M. (2003). Factors affecting biofilm accumulation in model distribution systems. *Journal of American Water Works Association.* 95(1): 87-97.

100. Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A. and Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology*.

101. Percival, S. L., Knapp, J. S., Wales, D. S and Edyvean, R. G. J. (1999). The effect of turbulent flow and surface roughness on biofilm formation in drinking water. *Journal of Industrial Microbiology and Biotechnology.* 22: 152-159.

102. Pintar, K. D. M. and Slawson, R. M. (2003). Effect of temperature and disinfection strategies on ammonia-oxidising bacteria in a bench-scale drinking water distribution system. *Water Research.* 37: 1805-1817.

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

104. Pinto, A. and Raskin, L. (2012b). PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. *PLoS.* 7(8): e43093.

105. 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.

106. Pinto, A., Marcus, D. N., Ijaz, U. Z., Bautista-de los Santos, Q. M., Dick, G. J. and Raskin, L. (2016). Metagenomic evidence for the presence of comammox *Nitrospira*-like bacteria in drinking water system. *mSphere.* 1(1): e00054-15.

107. Prest, E. I., Hammes, F., Kötzsch, S., Van Loosdrecht, M. C. M. and Vrouwenvelder, J. S. (2013). Monitoring microbiological changes in drinking water systems using a fast and reproducible flow cytometric method. *Water Research*. 47(19): 7131-7142.

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

109. Prest, E. I., Weissbrodt, D. G., Hammes, F., van Loosdrecht, M. C. M. and Vrouwenvelder, J. S. (2016b). Long-term bacterial dynamics in a full-scale drinking water distribution system. *PloS one.* 11(10): e0164445.

110. Ramos, H. M., Loureiro, D., Lopes, A., Fernandes, C., Covas, D., Reis, L. F. and Cunha, M. C. (2010). Evaluation of chlorine decay in drinking water systems for different flow conditions: from theory to practice. *Water Resources Management.* 24(4): 815-834.

111. Regan, J. M., Harrington, G. W and Noguera, D. R. (2002). Ammonia- and nitrite-oxidizing bacterial communities in a pilot-Scale chloraminated drinking water distribution system. *Applied and Environmental Microbiology.* 68(1): 73-81.

112. Regan, J. M., Harrington, G. W., Baribeau, H., De Leon, R. and Noguera, D. R. (2003). Diversity of nitrifying bacteria in full-scale chloraminated distribution systems. *Water Research.* 37(1): 197-205.

113. Rittmann, B. E. and Snoeyink, V. L. (1984). Achieving biologically stable drinking water. *Journal -American Water Works Association.* 76(10): 106-114.

114. 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.

115. Rossman, L. A., Clark, R. M. and Grayman, W. M. (1994). Modeling chlorine residuals in drinking-water distribution systems. *Journal of Environmental Engineering.* 120(4): 803-820.

116. Salipante, S. J., Kawashima, T., Rosenthal, C., Hoogestraat, D. R., Cummings, L. A., Sengupta, D. J., Harkins, T. T., Cookson, B. T. and Hoffman, N. G. (2014). Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. *Applied and Environmental Microbiology.* AEM-02206.

117. Sathasivan, A., Fisher, I. and Tam, T. (2008). Onset of severe nitrification in mildly nitrifying chloraminated bulk waters and its relation to biostability. *Water Research.* 42(14): 3623-3632.

118. Sawade, E., Monis, P., Cook, D. and Drikas, M. (2016). Is nitrification the only cause of microbiologically induced chloramine decay? *Water Research.* 88:904-911.

119. Schmeisser, C., Stockigt, C., Raasch, C., Wingender, J., Timmis, K. N., Wenderoth, D. F., Flemming, H. C., Liesegang, H., Schmitz, R. A., Jaeger, K. E. and Streit, W. R. (2003). Metagenome survey of biofilms in drinking water networks. *Applied and Environmental Microbiology.* 69(12): 7298-7309.

120. Servais, P., Anzil, A. and Ventresque, C. (1989). Simple method for determination of biodegradable dissolved organic carbon in water. *Applied and Environmental Microbiology.* 55(10): 2732-2734.

121. Shade, A., Jones, S. E., Caporaso, J. G., Handelsman, J., Knight, R., Fierer, N. and Gilbert, J. A. (2014). Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *mBio ASM.* 5(4): 1-9.

122. Sharpton, T.J. (2014). An introduction to the analysis of shotgun metagenomic data. *Frontiers in Plant Science.* 5: 209.

123. Siqueira, V. M. and Lima, N. (2013). Biofilm formation by filamentous fungi recovered from a water system. *Journal of Mycology.* 1-9.

124. Spang, A., Hatzenpichler, R., Brochier-Armanet, C., Rattei, T., Tischler, P., Spieck, E., Streit, W., Stahl, D. A., Wagner, M and Schleper, C. (2010). Distinct gene set in two different lineages of ammonia-oxidising archaea supports the phylum Thaumarchaeota. *Trends in Microbiology.* 18: 331-340.

125. Srinivasan, S., Harrington, G. W., Xagoraraki, I. and Goel, R. (2008). Factors affecting bulk to total bacteria ratio in drinking water distribution systems. *Water Research.* 42(13): 3393-3404.

126. Stahl, D. A. and de la Torre, J. R. (2012). Physiology and diversity of Ammonia-oxidising archaea. *Annual Review of Microbiology.* 66: 83-101.

127. Sun, H., Shi, B., Bai, Y. and Wand, D. (2014). Bacterial community of biofilms developed under different water supply conditions in a distribution system. *Science of The Total Environment.* 472: 99-107.

128. Szabó, K. É., Itor, P. O., Bertilsson, S., Tranvik, L. and Eiler, A. (2007). Importance of rare and abundant populations for the structure and functional potential of freshwater bacterial communities. *Aquatic Microbial Ecology.* 47(1): 1-10.

129. Szewzyk, U., Szewzyk, R., Manz, W. and Schleifer, K. H. (2000). Microbiological safety of drinking water. *Annual Reviews in Microbiology.* 54(1): 81-127.

130. Tatari, K., Musovic, S., Gülay, A., Dechesne, A., Albrechtsen, H. J. and Smets, B. F., (2017). Density and distribution of nitrifying guilds in rapid sand filters for drinking water production: Dominance of Nitrospira spp. *Water Research*. 127: 239-248.

131. Thomas, J. M. and Ashbolt, N. J. (2011). Do Free-Living amoebae in treated drinking water systems present an emerging health risk? *Environmental Science and Technology*. 45(3): 860-869.

132. Thomas, T., Gilbert, J. and Meyer, F. (2012). Metagenomics-a guide from sampling to data analysis. *Microbial Informatics and Experimentation*. 2(1): 3.

133. Tsai, Y. (2006). Interaction of chlorine concentration and shear stress on chlorine concentration, biofilm growth and particulate matter. *Bioresorce Technology*. 97: 1912-1919.

134. Tuomisto, H. (2010). A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*. 33: 2-22.

135. Van Der Kooij, D., Visser, A. and Hijnen, W. A. M. (1982). Determining the concentration of easily assimilable organic carbon in drinking water. *Journal of American Water Works Association*. 74(10): 540-545.

136. Van der Kooij, D. (1992). Assimilable organic carbon as an indicator of bacterial regrowth. *Journal of American Water Works Association*. 84(2): 57-65.

137. Van der Wende, E., Characklis, W. G. and Smith, D. B. (1989). Biofilms and bacterial drinking water quality. *Water Research*. 23(10): 1313-1322.

138. Van der Wielen, P. W. J. J., Voost, S. and van der Kooij, D. (2009). Ammonia-oxidising bacteria and archaea in groundwater treatment and drinking water distribution systems. *Applied and Environmental Microbiology*. 75(14): 4687-4695.

139. Van Kessel, M. A., Speth, D. R., Albertsen, M., Nielsen, P. H., den Camp, H. J. O., Kartal, B., Jetten, M. S. and Lücker, S. (2015). Complete nitrification by a single microorganism. *Nature*. 528(7583): 555.

140. Vasconcelos, J. L., Rossman, L. A., Grayman, W. M., Boulos, P. F. and Clark, R. M. (1997). Kinetics of chlorine decay. *Journal of American Water Works Association*. 89(7): 54-65.

141. Vikesland, P. J., Ozekin, K. and Valentine, R. L. (2001). Monochloramine decay in model and distribution system waters. *Water Research*. 35(7): 1766-1776.

142. Volk, C. J. and LeChevallier, M. W. (1999). Impacts of the reduction of nutrient levels on bacterial water quality in distribution systems. *Applied and Environmental Microbiology.* 65(11): 4957-4966.

143. Von Gunten, U. (2003). Ozonation of drinking water: Part II. Disinfection and by-product formation in presence of bromide, iodide or chlorine. *Water Research.* 37: 1469-1487.

144. Wang, H., Edwards, M., Falkinham, J. O. and Pruden, A. (2012). Molecular survey of the occurrence of *Legionella* spp., *Mycobacterium* spp., *Pseudomonas aeruginosa*, and amoeba hosts in two chloraminated drinking water distribution systems. *Applied and Environmental Microbiology.* 78(17): 6285-6294

145. Wang, H., Pryor, M. A., Edwards, M. A., Falkinham III, J. O. and Pruden, A. (2013). Effect of GAC pre-treatment and disinfectant on microbial community structure and opportunistic pathogen occurrence. *Water Research.* 47(15): 5760-5772.

146. Wang, H., Masters, S., Edwards, M. A, Falkinham, J. O. and Pruden, A. (2014). Effect of disinfectant, water age and pipe material on bacterial and eukaryotic community structure in drinking water biofilm. *Envrionmental Science and Technology.* 48: 1426-1435.

147. Wang, Y., Ma, L., Mao, Y., Jiang, X., Xia, Y., Yu, K., Li, B. and Zhang, T. (2017). Comammox in drinking water systems. *Water Research.* 116: 332-341.

148. World Health Organization (2011). Guidelines for drinking-water quality: World Health Organization. *Distribution and Sales, Geneva.* 27.

149. 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 of American Water Works Association.* 88(7): 74-84.

150. Wingender, J. and Flemming, H. C. (2004). Contamination potential of drinking water distribution network biofilms. *Water Science and Technology.* 49(11): 277-286.

151. Wingender, J. and Flemming, H. C. (2011). Biofilms in drinking water and their roles as reservoir for pathogens. *International Journal of Hygiene and Environmental Health.* 214(6): 417-423.

152. Wolfe, R. L., Lieu, N. I., Izaguirre, G. and Means, E. G. (1990). Ammonia-oxidising bacteria in a chloraminated distribution system: seasonal occurrence, distribution and disinfectant resistance. *Applied and Environmental Microbiology.* 56(2): 451-462.

153. Wu, H., Zhang, J., Mi, Z., Xie, S., Chen, C. and Zhang, X. (2015). Biofilm bacterial communities in urban drinking water distribution systems transporting waters with different purification strategies. *Environmental Biotechnology.* 99: 1947-1955.

154. You, J., Das, A., Dolan, E. M. and Hu, Z. (2009). Ammonia-oxidising archaea involved in nitrogen removal. *Water Research.* 43: 1801-1809.

155. Yu, J., Kim, D. and Lee, T. (2010). Microbial diversity in biofilms on water distribution pipes of different materials. *Water Science and Technology.* 61(1): 163-171.

156. Zhang, Y., Love, N. and Edwards, M. (2009). Nitrification in drinking water systems. *Critical Reviews in Environmental Science and Technology*. 39: 153-208.

157. Zhang, M., Liu, W., Nie, X., Li, C., Gu, J. and Zhang, C. (2012). Molecular analysis of bacterial communities in biofilms of a drinking water clearwell. *Microbes Environment.* 27(4): 443-448.

158. Zlatanović, L., Van Der Hoek, J. P. and Vreeburg, J. H. G. (2017). An experimental study on the influence of water stagnation and temperature change on water quality in a full-scale domestic drinking water system. *Water Research*. 123: 761-772.

CHAPTER 2

LONG-TERM SPATIAL AND TEMPORAL MICROBIAL COMMUNITY DYNAMICS IN A LARGE-SCALE DRINKING WATER DISTRIBUTION SYSTEM WITH MULTIPLE DISINFECTANT REGIMES

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.

Author Contributions:

SC Potgieter performed all sampling and consequent analyses. SC Potgieter was involved in the conceptualisation and design of the research and the interpretation of the data as well as the drafting of the manuscript. AJ Pinto was involved in the conceptualisation and design of the research and provided guidance and expertise with regards to sampling processes, data analysis and interpretation as well as the drafting of the manuscript. M Sigudu, H Du Preez and E Ncube were part of the Rand Water team involved in this project. M Sigudu and H Du Preez were involved in the conceptualisation and design of the project and provided logistical assistance with regards to sampling. SN Venter was involved in conceptualisation and design of the research and the interpretation of the data as well as the drafting of the manuscript.

Abstract

Long-term spatial-temporal investigations of microbial dynamics in full-scale drinking water distribution systems are scarce. These investigations can reveal the process, infrastructure, and environmental factors that influence the microbial community, offering opportunities to re-think microbial management in drinking water systems. Often, these insights are missed or are unreliable in short-term studies, which are impacted by stochastic variabilities inherent to large full-scale systems. In this two-year study, we investigated the spatial and temporal dynamics of the microbial community in a large, full scale South African drinking water distribution system that uses three successive disinfection strategies (i.e. chlorination, chloramination and hypochlorination). Monthly bulk water samples were collected from the outlet of the treatment plant and from 17 points in the distribution system spanning nearly 150 kilometres and the bacterial community composition was characterised by Illumina MiSeq sequencing of the V4 hypervariable region of the 16S rRNA gene. Like previous studies, *Alpha-* and *Betaproteobacteria* dominated the drinking water bacterial communities, with an increase in *Betaproteobacteria* post-chloramination. In contrast with previous reports, the observed richness, diversity, and evenness of the bacterial communities were higher in the winter months as opposed to the summer months in this study. In addition to temperature effects, the seasonal variations were also likely to be influenced by changes in average water age in the distribution system and corresponding changes in disinfectant residual concentrations. Spatial dynamics of the bacterial communities indicated distance decay, with bacterial communities becoming increasingly dissimilar with increasing distance between sampling locations. These spatial effects dampened the temporal changes in bulk water community and were the dominant factor when considering the entire distribution system. However, temporal variations were consistently stronger as compared to spatial changes at an individual sampling location and demonstrated seasonality. This study emphasises the need for long-term studies to comprehensively understand the temporal patterns that would otherwise be missed in short-term investigations. Furthermore, systematic long-term investigations are particularly critical towards determining the impact of changes in source water quality, environmental conditions, and process operations on the changes in microbial community composition in drinking water distribution systems.

1. Introduction

Drinking water distribution systems (DWDSs) are designed and maintained to transport chemically and biologically safe, potable water to consumers. These systems are complex aquatic environments with multiple ecological niches that support microbial growth through the different stages of the DWDS. The microbial ecology of DWDSs is governed by environmental and engineering factors as well as operational conditions that influence the composition and structure of bacterial communities present in the biofilms, bulk water and sediments (Liu *et al.,* 2013; Prest *et al*., 2016a; Liu *et al*., 2017). Despite disinfection during water treatment, microorganisms grow during distribution with reported microbial cell numbers ranging between $10^4 – 10^6$ cells per litre (Hammes *et al*., 2008). This persistent microbial community can be highly diverse including bacteria, archaea, free living amoebae, fungi and viruses (You *et al*., 2009; Thomas and Ashbolt, 2011; Siqueira and Lima, 2013; Liu *et al.,* 2013; Gall *et al*., 2015).

The concentration and composition of microorganisms within DWDSs is influenced by multiple treatment processes, specifically primary and secondary disinfection through the use of chlorine and/or chloramine dosing, respectively (Gomez-Alvarez *et al*., 2012). This final step of drinking water treatment profoundly alters the DWDS microbiome structure and composition and significantly reduces bacterial cell numbers depending on the disinfectant used (Proctor and Hammes, 2015; Prest *et al*., 2016a). However, despite disinfection, DWDS microbial communities persist in a limited, low-nutrient environment and disinfection may even select for unwanted bacteria such as *Mycobacteria*, ammonia- and nitrite-oxidising bacteria (Proctor and Hammes, 2015).

The bulk water is the primary medium for the spread of microorganisms, nutrients, and particles throughout the DWDS and it feeds into the point-of-use (PoU), which is the final point of consumer exposure to the drinking water microbiome (Liu *et al.,* 2013; Bautista-de los Santos *et al*., 2016). It has long been assumed that bacteria in the bulk water originate from detachment of biofilms or re-suspension of the sediments rather than bacterial growth in the bulk phase itself (Prest *et al*., 2016a). However, microbial communities within biofilms and bulk water have been shown to be distinct and biofilms from DWDS pipe walls may only have a minor impact on the microbial community in the bulk water at the point of consumption (Henne *et al*., 2012; Liu *et al.,* 2014). Bulk water communities have been found

to be spatially stable over short time scales irrespective of the DWDS sample location (Lautenschlager *et al*., 2013; Roeselers *et al.,* 2015) and have also been reported to display annually reproducible temporal trends (Pinto *et al*., 2014).

Due to developments in high throughput sequencing, our understanding of the DWDS microbiome has significantly improved (Proctor and Hammes, 2015; Bautista-de los Santos *et al*., 2016). Several studies have highlighted the effects of specific characteristics on the dynamics of microbial communities, including different treatment strategies (Gomez-Alvares *et al*., 2012; Hwang *et al*., 2012; Wang *et al*., 2014b, Bautista-de los Santos *et al*., 2016), distribution (Nescerecka *et al*., 2014; Shaw *et al*., 2015), process operations (Pinto *et al*., 2012; Lautenschlager *et al*., 2014; Wu *et al*., 2015), hydraulic conditions (Douterelo *et al*., 2014), water age and residence time (Wang *et al*., 2014a; Prest *et al.,* 2016a; Zlatanovic *et al*., 2017) as well as pipe material (Niquette *et al*., 2000; Wang *et al*., 2014a).

Although seasonal changes in DWDS microbial communities have previously been investigated (McCoy and VanBriesen, 2012 and 2014; Pinto *et al*., 2014; Ling *et al*., 2016; Zlatanovic *et al*., 2017), long-term, in-depth investigations of spatial and temporal dynamics of DWDS microbial communities are rare. Temporal dynamics of DWDSs cannot be accurately described without an extensive, long-term and high-frequency sampling strategy (Prest *et al*., 2016b). Such long-term investigations can provide insight into robust processes, infrastructure, and environmental factors (i.e. temperature, pH, disinfectant residuals, turbidity, etc.) that influence the microbial community, presenting opportunities to re-think the management of microbial growth and community composition in drinking water systems. Often, these insights into seasonal variations are unreliable in short-term or low sampling frequency studies due to the stochastic variabilities inherent to large full-scale systems. Short term or single time point sampling cannot be extrapolated to represent other times of the year, as several studies have shown bacterial communities can undergo significant temporal variations even within a single year (Pinto *et al*., 2014; Prest *et al* 2016b).

The primary challenge in terms of defining the drinking water microbiome is that it changes dynamically through all stages of the DWDS. The DWDS represents a microbial continuum, where the given volume of water and associated microbial community migrates while being influenced by changing disinfectant residual concentrations, nutrient bio-availability, and by

the microbial communities in the biofilms and sediments. Therefore, an integral component of this study was to assess how spatial-temporal variation shapes the drinking water microbial community as it traverses through varying disinfection regimes (Proctor and Hammes, 2015). The current study aims to understand how the temporal and spatial dynamics of a unique and complex large-scale drinking water distribution network shapes the bacterial community structure and composition. To this end, a two year sampling campaign was conducted for a complex and multiple branching section of the DWDS that encompasses a three-stage disinfectant strategy i.e., initial chlorine dosing followed by the addition of chloramine and lastly, hypochlorite. The objectives were to: (i) assess long-term seasonal variations in the bacterial community of the DWDS over two years, (ii) determine the effects of the spatial configuration of DWDS on the bacterial community, considering the use of three different disinfectant residuals, and (iii) to understand the interplay between the temporal and spatial dynamics within the DWDS as a whole as well as within each disinfection strategy. This long-term study aims to provide a unique insight into physical-chemical factors impacting the spatial-temporal dynamics of the drinking water microbiome in the DWDS with multiple disinfectant regimes. This knowledge combined with previous (and future) spatial-temporal studies may help water utilities identify strategies to manage the drinking water microbiome in the DWDS to ensure its safety and stability.

2. Methodology

2.1 Site description and sample collection

Sampling was conducted at the outlet of a South African full-scale drinking water treatment plant (DWTP) and corresponding DWDS, from October 2014 to September 2016. In its entirety, this water utility serves consumers over a vast network, stretching over 3056 kilometres covering 18,000 km^2. The DWDS feeds 58 service reservoirs which supply large metropolitan and local municipalities as well as mines and industries with on average, 3653 million litres of water supplied daily to approximately 11 million people. Due to the complexity, multiple distribution branches and vastness of this DWDS, a section of the distribution system was selected for this study, originating from the DWTP and spanning approximately 150 km of the corresponding DWDS pipeline.

Treatment of source water derived from surface water consists of coagulation with polymeric coagulants, flocculation and addition of lime (55-70 mg/L calcium hydroxide),

sedimentation, pH adjustment with CO_2 gas followed by filtration (rapid gravity sand filters) and finally disinfection, which includes 3 disinfection strategies. First, the filter effluent is dosed with chlorine where liquid chlorine is evaporated and bubbled into carriage water to be dosed into the main water for disinfection. Chlorine dosages vary depending on the source water quality and the system demand, which is typically higher in the summer months. The total residual chlorine at these dosages varies between 1 mg/L in summer and 1.5 mg/L in winter after 20 minute contact time. Second, within this selected section of the DWDS, chlorinated water leaving the DWTP is again dosed with chloramine (0.8 to 1.5 mg/L) at a secondary disinfection boosting station approximately 23 km from the treatment plant. Here, monochloramine residuals vary on average between 0.8 mg/L in the autumn and 1.5mg/L in the spring. And finally, bulk water is again disinfected with hypochlorite at locations towards the end of the sampled DWDS section (approximately 120 km from the DWTP). In this hypochlorinated section of the DWDS, total residual chlorine varies on average between 0.6 mg/L in the summer and 1.2 mg/L in the winter. Free residual chlorine remains constant both temporally and spatially, varying on average between 0.3 and 0.2 mg/L. The dissolved organic carbon (DOC) concentrations in the DWDS were relatively constant throughout the duration of the study (i.e. between 3.5 and 4.2 mg Carbon/L, in autumn and spring, respectively). Further details on range of physical-chemical parameters, length, composition, and age of the pipe line sections connecting the DWDS sample locations were obtained from the utility for this study (Figure 1 and Table A4).

Bulk water samples were collected from 18 locations including the outlet of the DWTP, immediately before the chlorinated water enters the DWDS, and 17 locations within the DWDS (including bulk water samples from pipeline and reservoirs). This included 2 chlorinated, 13 chloraminated, and 3 hypochlorinated bulk water samples. Sampling occurred consecutively for 2 days on a monthly basis for 2 years, except for January 2016 and sample in July 2016 for sample site CHM2.2, resulting in the collection of 413 samples. Sample site descriptions and DWDS layout are provided in Figure 1 and Table 1.

2.2 Sample processing

Bulk water samples were collected in 8L sterile Nalgene polycarbonate bottles and transported to the laboratory cold where they were kept at 4°C for 24 to 48 hours until filtered. Samples were filtered to harvest microbial cells by pumping the collected bulk water

through STERIVEX™ GP 0.22 µm filter units (Millipore) using a Gilson® minipuls 3 peristaltic pump. Typically, 8L of bulk water were filtered for each sample. For samples collected directly after a disinfection, 16L of bulk water was filtered. The filters were kept in the dark and stored at -20°C until processing and DNA extraction. A traditional phenol/chloroform extraction method optimised by Pinto *et al.* (2012) was used for the isolation of DNA from cells immobilised on filter membranes. This protocol represents a modified version of the protocol described by Urakawa *et al.* (2010). Extracted DNA was sent to the Department of Microbiology and Immunology at the University of Michigan, Ann Arbor, USA for sequencing of the V4 hypervariable region of the 16S rRNA gene using the Illumina MiSeq platform. The dual-index paired-end sequencing approach, described by Kozich *et al.* (2013), resulted in paired reads with each read pair with a length of 250 nucleotides. All raw sequence data have been deposited with links to BioProject accession number PRJNA445682 in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/).

2.3 Sequencing data processing

Due to unsuccessful sequencing attempts (i.e. failed DNA amplification), 65 of the 413 samples collected were excluded (excluded samples are described in Table A1) and the number of samples per sample site is indicated in Table 1. Failure of DNA amplification may have been a consequence of multiple factors i.e. low DNA concentrations, the potential presence of inhibitors (i.e. humic substances, phenolic compounds) and/or extensive DNA damage caused by high levels of disinfectant (i.e. chlorine) (Van Aken and Lin, 2011; Schrader *et al.*, 2012). Sequence processing and analysis of the remaining 348 samples was performed using mothur (version 1.35.1) (Schloss *et al.*, 2009) according to the protocol outlined previously (Kozich, *et al.*, 2013). Merging of the forward and reverse reads yielded 11,568,699 sequences and resulting sequences were screened by allowing a maximum length of 275 base pairs (bp) and minimum length of 250 bp. Sequences with more than eight homopolymers and any ambiguities were removed. Sequences were aligned to the SILVA reference database (Quast *et al.*, 2013) and resulting alignments were trimmed using the filter.seqs option in mothur, ensuring that all sequences were aligned along similar regions of the V4 region of the 16S rRNA gene. The filtered and aligned sequences were further processed through the pre.cluster option by using a pseudo-single linkage algorithm with a 2-bp similarity threshold. Chimeras were identified using UCHIME (Edgar *et al.*, 2011) and

removed. The remaining 8,568,237 quality filtered sequences were classified using the Greengenes database (DeSantis *et al*., 2006), with a threshold confidence level of 80%. Sequences with an unknown domain level of taxonomy were discarded, resulting in a total of 8,314,324 sequences with an average of 23,892 ± 13,260 sequences per sample and the minimum and maximum number of sequences per sample being 1007 and 71843, respectively. Sequences were aligned using average neighbour algorithm into operational taxonomic units (OTUs) with a similarity cut-off of 97%.

2.4 Statistical analysis

Multivariate Cut-off Level Analysis, MultiCoLA (Gobet *et al*., 2010) was applied to evaluate the extent to which each OTU contributes to the structure of each community and filter the dataset to only retain OTUs that explained majority of the community structure variability. Specifically, the dataset was sorted according to the decreasing total sum of OTU sequences and then the top 1% of OTUs were retained, where each of the top 1% OTUs retained had a minimum of 3194 sequences. Mantel's test, was performed in R using the mantel function in the vegan package (Oksanen *et al*., 2015), between the structure based dissimilarity matrix constructed using the original dataset and one constructed using only the dataset consisting of the retained top 1% OTU dataset to determine whether the variability within the bacterial community structure was maintained within the smaller subset of OTUs.

Alpha diversity indices (observed richness, Shannon Diversity Index, Inverse Simpson Diversity Index and Pielou's evenness) were calculated using the summary.single function in mothur (Schloss *et al*., 2009) incorporating the parameters, iters=1000 and subsampling=1007 (sample containing the least number of sequences). Good's coverage estimates were included to determine the percentage of coverage associated with each sample after subsampling. Testing for normality using Shapiro-Wilk test and Q-Q plot, and Leven's test for homogeneity of variance revealed that alpha-diversity indices had a non-normal distribution, using the stats (R Core Team, 2015) and car (Fox and Weisberg, 2011) packages, respectively. Kruskal-Wallis one-way analysis of variance and post-hoc Dunn's test were performed in R using the stats and dunn.test (Dinno, 2017) packages, respectively to determine whether alpha-diversity indices were significantly different when grouped based on DWDS sample location, month, season or disinfectant used.

Beta diversity analyses were performed to compare samples using OTU-level assignment (Jaccard and Bray-Curtis) and phylogenetic placement (weighted and unweighted UniFrac). Membership (Jaccard and unweighted UniFrac) and structure based (Bray-Curtis and weighted UniFrac) beta-diversity metrics were calculated using mothur (Schloss *et al*., 2009). Phylogenetic-based metrics were obtained by constructing a phylogenetic tree containing all sequences (97% similarity threshold), using the clearcut command in mothur (Evans *et al*., 2006; Lozupone *et al.,* 2011). All matrices were calculated after 1000 subsamplings of the entire data set (iters=1000) to the number of the least number of sequences (n = 1007) ensuring that all samples were compared with the same sequence depth. All four beta-diversity metrics and metadata files containing sample location, disinfection type, seasons and months were imported to R (http://www.R-project.org) for statistical analyses, including permutational analysis of variance (PERMANOVA) using the adonis function in the vegan package (Oksanen *et al*., 2015). Analysis of Molecular Variance (AMOVA) was performed in mothur, using the amova function, to determine the effect of different groupings of samples based on DWDS sample location, month, season and disinfection type (Excoffier, 1993; Anderson, 2001). Principal-coordinate analysis (PCoA) using Bray-Curtis and Jaccard distances were performed using the phyloseq package (McMurdie and Holmes, 2013). All plots were constructed using the ggplot2 package (Wickham, 2009).

3. Results

3.1 Small percentage of OTUs maintain the majority of the spatial and temporal trends

The bacterial community within all samples was taxonomically diverse with 9,516 OTUs being identified across all samples. Of the 99% low-abundance OTUs removed using MultiCoLA (Gobet *et al.,* 2010), 69.5% had total reads of 10 or less. Furthermore, of those OTUs removed, 14.4% were doubletons and 26.1% were singletons. The 95 OTUs (Table A2) that were retained constituted >1% of the total sequence counts and were shared among all sample points. These frequently detected 95 OTUs made up for 90% of the total sequences post quality filtering. Furthermore, these 95 OTUs captured 99% of spatial-temporal variability between samples (Mantel's $R_{Bray-Curtis}$ = 0.993; p = 0.001). We also assessed the effect of subsampling, employed as means of normalizing variability in library size across samples, on the diversity captured within each sample based on Good's coverage analyses.

This indicated that subsampling at a library size of 1007 sequences captured the majority of the richness for all samples (i.e., Good's coverage = 96.6 ± 0.6%).

3.2 Dominance of *Alpha-* and *Betaproteobacteria* varies depending on the type of disinfectant residual

The drinking water microbial community was dominated by bacteria (99.7% sequences were of bacterial origin), with archaea constituting only 0.3% of the total sequences. Of the 60 bacterial phyla identified, *Proteobacteria* was the most dominant across all samples with a mean relative abundance (MRA) of 78.2 ± 12.4%, constituting 82.4% of the total sequences (6,861,465 sequences) and 48 to 89% of the bacterial community in any given sample. The second most dominant phylum was *Planctomycetes* with a MRA of 10.1 ± 8.6%, constituting 7.7% of the total sequences (Table A3).

Further characterisation of the bacterial classes revealed *Alphaproteobacteria*, *Betaproteobacteria*, *Planctomycetia* and *Gammaproteobacteria* dominated, with MRAs of 49.6 ± 8.3%, 22.6 ± 12.2%, 8.8 ± 7.2% and 4.5 ± 2.4% across all sample locations, respectively (Figure 2). *Betaproteobacteria* showed a significant increase in relative abundance following chloramination (MRA of 2.6 ± 3.1% in CHM1 to MRA of 37.8 ± 17.2% in CHM2). More specifically, members belonging to *Betaproteobacteria* (i.e. OTUs classified to the genus *Nitrosomonas* and the family *Nitrosomonadaceae*, the genus *Gallionella*) as well as members from the genus *Nitrospira*, increased after chloramination. However, some OTUs persisted after chloramination, including members of the phyla *Planctomycetes*, *Cyanobacteria* and the proteobacterial class *Alphaproteobacteria* (e.g. members of the genus *Hyphomicrobium*). *Alphaproteobacteria* remained relatively stable throughout the duration of the study, but showed a higher abundance in the summer months (MRA 52.9 ± 21.0%) compared to the winter months (MRA 44.7 ± 17.7%). Similarly, the relative abundance of *Betaproteobacteria* was highest in the summer months particularly within the chloraminated section of the DWDS (MRA 32.0 ± 21.0%) with a maximum MRA of 41.9 ± 25.4% in December 2015 but decreased in the spring (MRA 18.1 ± 17.8%). Conversely, the classes *Planctomycetia* and *Gammaproteobacteria* showed a decreased relative abundance in the summer months (MRA 5.4 ± 10.6% and MRA 1.7 ± 4.3%, respectively) to maximum relative abundance in the spring months (MRA 13.7 ± 12.5% and MRA 7.2 ± 13.5%, respectively). *Planctomycetia* also showed a decrease in relative

abundance following chloramination and increased residence time in the reservoirs (CHM2) from a MRA of $22.4 \pm 6.6\%$ in samples before the reservoirs at CHM2 to a MRA of $6.5 \pm 2.0\%$ after the reservoirs. This decrease in *Planctomycetia* strongly correlated with an increase in *Betaproteobacteria* relative abundance (Pearson's $R = -0.74$, $p < 0.001$). Other bacterial classes with a MRA above 1% across all sample sites included the *Cyanobacterial* class 4C0d-2, *Nitrospira, Actinobacteria* and *Phycisphaerae* (MRA $2.0 \pm 1.9\%$, $1.6 \pm 2.0\%$, $1.4 \pm 1.3\%$ and $1.2 \pm 1.6\%$, respectively). The remaining 140 classes and unclassified taxa constituted 9.9% of the total sequences with varying relative abundance throughout the year.

At the OTU level, the most abundant OTUs (overall contribution to abundance >1%) are shown in Table 2. The most abundant OTU was classified as *Nitrosomonas* (class: *Betaproteobacteria*, family: *Nitrosomonadaceae*) with MRA of $15.0 \pm 18.9\%$ and constituted 18.5% of the total sequences and had 100% sequence similarity to *Nitrosomonas oligotropha*. A significant increase in the relative abundance of *N. oligotropha*-like OTU was observed following chloramination and increased residence time in the reservoirs (from MRA $0.2 \pm 0.4\%$ in CHM1 to MRA $18.8 \pm 22.1\%$ in CHM2). The relative abundance of OTUs classified as *Nitrosomonas* increased from 4-11% of the total betaproteobacterial sequences in chlorinated water (MRA $0.3 \pm 0.8\%$) to 23-40% of betaproteobacterial sequences post chloramination (MRA $16.2 \pm 19.3\%$). The relative abundance of four most abundant OTUs (55.4% of the total sequences) (Table 2) was strongly correlated with that of OTU 1 (Genus: *Nitrosomonas*) and OTU 2 (Order: *Rhizobiales*), which increased in relative abundance in the summer months, whereas OTU 3 (Order: *Rhizobiales*) and OTU 4 (Genus: *Sphingomonas*) increased in relative abundance in the winter months. Specifically, *N. oligotropha*-like OTU reached its maximum abundance in the summer months (Dec - Feb) (MRA $25.3 \pm 22.0\%$ in summer compared to MRA $5.3 \pm 11.5\%$ in winter).

3.3 Increased richness in winter months with seasonal cycling

Alpha-diversity measures showed strong seasonal trends over the two years of this study. Richness (observed number of OTUs) was found to be higher in the winter months (July and August, average observed OTUs of 293 ± 161) compared to the summer months (December – February, average observed OTUs 207 ± 92) (Figure 3A). In the winter months, bacterial communities were also more diverse (average Shannon Diversity Index winter: 3.00 ± 0.69 vs summer: 2.00 ± 0.70; average Inverse Simpson Diversity Index winter: 10.04 ± 6.20 vs.

summer: 5.26 ± 4.39) and even (average Pielou's evenness winter: 0.54 ± 0.12 vs. summer: 0.39 ± 0.13) (Figure A1).

Significant seasonal differences in the richness, diversity, and evenness of the bacterial community were observed (Kruskal-Wallis, $p < 0.05$). Post-hoc Dunn's test revealed significant differences between the summer and winter months (Bonferroni-corrected, $p < 0.05$ for all alpha-diversity measures). The changes in alpha diversity measures correlated with seasonal temperature changes as well as varying chlorine (i.e. total and free chlorine) and monochloramine residuals within the DWDS. The increase in richness in the winter months showed a moderate negative correlation with water temperature (Pearson's R = -0.56, p <0.001) (Figure 3B). Similar correlations were observed for Shannon diversity (Pearson's R = -0.55, p <0.001), Inverse Simpson diversity (Pearson's R = -0.46, p <0.001) and Pielou's evenness (Pearson's R = -0.47, p <0.001). Conversely, richness showed a moderate positive correlation with total chlorine (Pearson's R = 0.48, p <0.001) and monochloramine (Pearson's R = 0.48, p <0.001) (Figure 3B and Table A4).

Seasonal trends were observed in bacterial community membership (Jaccard and unweighted UniFrac) and structure (Bray-Curtis and weighted UniFrac) across all sample locations, indicating that samples 6-7 months apart showed increased dissimilarity followed by a decrease in dissimilarity 11-12 months apart. However, these changes in dissimilarities were marginal suggesting relative temporal stability. These seasonal trends were more clearly reflected within the bacterial community structure (Bray-Curtis and weighted UniFrac) at individual sample sites. Specifically, reservoirs (CHM2.1, CHM5.2 and CHM5.3) and pipeline (CHM3.1-3.3, CHM4.1 and CHM4.2) samples sites within the chloraminated section of the DWDS showed significant seasonal trends with an increase in dissimilarity between samples 6 months apart (Bray-Curtis: 0.78 ± 0.19, weighted UniFrac: 0.53 ± 0.16) and a decrease in dissimilarity in samples 12 months apart (Bray-Curtis: 0.68 ± 0.19, weighted UniFrac: 0.43 ± 0.17).

Principal coordinate analysis (PCoA) based on Bray-Curtis distances revealed some seasonal cycling in community structure (Figure 4B). PCoA analyses using membership based Jaccard distances showed similar results (Figure A2). Although clustering was not pronounced and some overlap between seasons was observed, seasons from the first year clustered with the

corresponding season of the following year, indicating seasonal cycling and suggested the potential for annual reproducibility in bacterial community membership and structure. More specifically, samples collected in summer and autumn clustered closer together, as well as those collected in spring and winter. These seasonal trends were more pronounced within the chloraminated (CHM) and hypochlorinated (HCHL) sections of the DWDS (Figure 4D and 4E, respectively). Though clear seasonal clustering was not observed for the chlorinated section of the DWDS, the data points clustered based on the year of sampling; more specifically samples from the second year clustered more closely together and distinct from the first years samples (Figure 4C). Differences were also observed in all beta-diversity metrics when samples were grouped based on the season in which they were collected, specifically between summer and winter (community membership [Jaccard and unweighted UniFrac: AMOVA, $F_{ST} \leq 3.98$, $p < 0.001$] and structure [Bray-Curtis and weighted UniFrac: AMOVA, $F_{ST} \leq 9.55$, $p < 0.001$]).

3.4 Increased temporal variation with increased distance from a disinfection point

The temporal variation was much more pronounced when considering individual sampling locations DWDS sample locations (Bray-Curtis, 0.72 ± 0.17 and weighted UniFrac, 0.52 ± 0.15) across the two year study period (Figure A3), as compared to the DWDS as a whole. On average, each sampling location showed 30-50% similarity in community structure between sampling time points. The bacterial community structure of sample locations immediately downstream of disinfection (i.e. CHL1, CHM1.1, CHM1.2 and HCHL1) showed minimal temporal variation. However, with increasing distance away from disinfection sites, particularly within the chloraminated section of the DWDS, samples sites showed increased temporal variation in community structure, particularly at the chloraminated reservoirs (CHM2 and CHM5) (Figure 5A and Figure A3). Similar trends in the temporal variation were also observed for community membership (Jaccard), however these were less pronounced (Figure 5B). The community membership was more dissimilar and the dissimilarity was less variable between temporal samples at each location (Jaccard; 0.78 ± 0.06 and unweighted UniFrac; 0.72 ± 0.05) with an average of only 20-30% similarity in community membership between sampling time points (Figure A3).

3.5 Significant distance decay in community structure with increasing distance between sample locations

Lower richness levels were observed for locations CHL1, CHL2 and CHM1 (observed OTUs: 137 ± 102, 101 ± 52 and 151 ± 52, respectively) (Figure A4), which corresponds to the location of these sample sites immediately after chlorination (CHL1) and immediately before and after chloramination (CHL2 and CHM1, respectively). Furthermore, these three locations showed significant differences in richness compared to the remaining DWDS sample locations (Dunn's test Bonferroni-corrected, $p < 0.05$). An increase in richness was observed within the chloraminated section of the DWDS, specifically following secondary disinfection at CHM1 (observed OTUs: 151 ± 52) and increased residence time in a reservoir at CHM4 (observed OTUs: 250 ± 106) (Figure A4). The highest levels of richness were observed in the chloraminated reservoirs at CHM5, with an average observed OTUs of 303 ± 127. Significant differences in the richness was observed when samples were grouped based on location (Kruskal-Wallis, $p < 0.05$). Further investigations revealed that these differences existed between locations before and after disinfection events, specifically after secondary disinfection with chloramine (Dunn's test Bonferroni-corrected $p < 0.05$ between CHM1 and CHM2), which correlated to the observed changes in richness. However this location based difference in richness was not observed for the other alpha diversity measures (Shannon Index, Inverse Simpson Index or Pielou's evenness, Kruskal-Wallis, $p > 0.05$).

Differences in microbial community structure were also observed following secondary disinfection with chloramine. PCoA ordination based on Bray-Curtis distances revealed limited spatial clustering within the community structure based on the different disinfection strategies (Figure 4A) with only chlorinated water samples (CHL1 and CHL2) clustered closer together. No clear clustering was observed for chloraminated or hypochlorinated samples. However, chlorinated samples were shown to be significantly different from samples containing the other two disinfectants (AMOVA, $F_{ST} \leq 9.18$, $p < 0.001$).

To assess the effect of distance between samples, the pairwise dissimilarity distances between individual samples were grouped based on the distance between them. The spatial dynamics of the bacterial community structure (Bray-Curtis and weighted UniFrac metrics) within the chloraminated section of the DWDS showed significant distance decay, with bacterial community structure becoming increasingly dissimilar with increasing distance between

sampling locations (R^2 = 0.14, p<0.001) (Figure 6). However, no significant distance decay features were observed for community membership (Jaccard and unweighted UniFrac).

3.6 Dissimilarity between DWDS sample locations conforms to the layout of the DWDS

Beta diversity distances between samples immediately after disinfection and all other samples within each disinfection section were performed in line with the layout of the DWDS. The two chlorinated samples (CHL1 and CHL2) were on average 67% dissimilar. However, within the chloraminated section, clear spatial differences in community structure (Bray-Curtis and weighted UniFrac) were observed with samples from location CHM1 and CHM2 showing an increase in dissimilarity. This correlated with the addition of chloramine at location CHM1 and increased residence in the reservoirs (CHM2.1 and CHM2.2). Dissimilarity in bacterial community structure decreased slightly in samples within chloraminated pipelines (locations CHM3 and CHM4) but again increased with chloraminated water entering the reservoirs at CHM5 (Figure 7). Samples within the hypochlorinated section showed decreased dissimilarity. This shows the variation in bacterial community structure (Bray-Curtis distances) as bulk water moves through the consecutive locations with an increase in dissimilarity at the reservoir sites. Similar variations in dissimilarity were observed with community membership (Jaccard and unweighted UniFrac), however the changes were marginal (Figure A5).

Furthermore, these variations in dissimilarity within each disinfection section correlate with changes in the disinfectant residual concentration, with both the chlorinated and chloraminated sections, demonstrating disinfectant decay. Within the chlorinated section the total chlorine decreased from 2.03 ± 0.14 mg/L in CHL1 to 0.97 ± 0.32 mg/L in CHL2. Within the chloraminated section both total chlorine and monochloramine concentrations decreased from location CHM1 to CHM5 (total residual chlorine: CHM1 2.20 ± 0.20 mg/L to CHM5 0.77 ± 0.62 mg/L; Monochloramine: CHM1 2.13 ± 0.30 mg/L to CHM5 0.66 ± 0.61 mg/L). Conversely, total residual chlorine concentrations remained relatively stable within the hypochlorinated section of the DWDS.

3.7 Temporal dynamics are dominant at a localised level

To further assess the impact of spatial and temporal dynamics on the bacterial community, samples were grouped based on season versus DWDS sample location and were compared using membership based (Jaccard and unweighted UniFrac) and structure based (Bray-Curtis and weighted UniFrac) metrics. PERMANOVA results (Table A5) of all four beta-diversity metrics showed that variations in the bacterial community over the DWDS as a whole, were best explained by sampling location (PERMANOVA, $R^2 \leq 0.21$) whereas, seasonal groupings had little or no impact on the DWDS bacterial community when considering all sample locations together (PERMANOVA, $R^2 \leq 0.07$).

Although the three disinfection sections did not cluster independently in the PCoA analyses (Figure 4A), when the bacterial communities were grouped based on the 3 disinfection strategies used and compared, the differences in community membership and structure were statistically significant (AMOVA, $F_{ST} \leq 10.92$, p < 0.001). This difference between PCoA and AMOVA analyses may emerge due to underlying temporal trends (Figure 4B). Specifically, the temporal variation in samples within each disinfection strategy may be larger than the differences between samples grouped across disinfection strategies. Therefore, in order to clearly understand the interplay between the spatial and temporal dynamics of the microbial community, each disinfection section was analysed separately [i.e. section 1: chlorinated water (CHL), section 2: chloraminated water (CHM) and section 3: hypochlorinated water (HCHL)] (Figure 1)]. PERMANOVA tests on these three defined sections based on all four beta diversity metrics revealed that although spatial groupings may best explain the overall variability among samples, temporal/seasonal groupings best explained the variability within each disinfection section. More specifically, this was clearly observed for the chlorinated section (section 1; CHL), where temporal groupings were better supported (PERMANOVA, yearly seasons: $R^2 \leq 0.41$ vs sample site: $R^2 \leq 0.05$) (Table A6) over spatial/location groupings. Spatial groupings in the chlorinated section (CHL) explained very little of the variation (approximately 5%) with low significance but considering that this section included only two sample locations (CHL1 and CHL2) this result is not surprising. Similarly, the hypochlorinated section (section 3; HCHL) included only the 3 hypochlorinated locations (HCHL1 - HCHL3), however these sample locations were more spatially variable which was reflected in the PERMANOVA results (yearly seasons: $R^2 \leq 0.22$ vs sample site: $R^2 \leq 0.12$) (Table A8). Although, temporal groupings explained more of

the variation in section 3 (HCHL), spatial groupings also had an effect on the bacterial community, specifically within community structure (Bray-Curtis and weighted UniFrac).

For the chloraminated section (section 2; CHM), the differences between temporal and spatial groupings were marginal and not well supported (Table A7). Temporal groupings explained more of the variation in bacterial community structure (PERMANOVA for Bray-Curtis and weighted UniFrac: $R^2 \leq 0.06$) with the effect of temporal groupings on community membership being insignificant. Conversely, spatial groupings explained slightly more of the variation in community membership (PERMANOVA, Jaccard and unweighted UniFrac $R^2 \leq 0.07$) and had no significant effect on community structure. However, the combination of both temporal and spatial groupings explained up to 36% of the variation, although only significant for community membership (Jaccard and unweighted UniFrac). These results indicate that although seasonal groupings may explain more of the variation within each section, spatial groupings may have more of an impact on the variation between samples in section 2 (CHM) than in section 1 (CHL) and 3 (HCHL). This observation could in part be due to the fact that section 2 includes more samples with greater distances between them and therefore allows the spatial dynamics in the DWDS to have more of an impact on the bacterial community.

4. Discussion

This study represents the first long-term spatial-temporal investigation of microbial community dynamics in a large DWDS that utilizes multiple disinfectant regimes (i.e. chlorination, chloramination and hypochlorination). In doing so, we provide unique insights into how a microbial community responds to different disinfectants and their concentrations as it migrates through the DWDS at multiple time-points over two years. Further, the duration of the study and the length of the DWDS studied allowed us for the first time to contextualise the importance of spatial compared to temporal variation. To our knowledge majority of spatial-temporal studies thus far have either been performed on smaller and/or for much shorter periods of time (McCoy and VanBriesen, 2012 and 2014; Pinto *et al*., 2014; Prest *et al*., 2016b; Zlatanovic *et al*., 2017).

4.1 Bacterial community composition

Consistent with previous studies, *Proteobacteria* dominated the bacterial community (Bautista-de los Santos *et al*., 2016). The phylum *Proteobacteria* and particularly the two classes *Alpha-* and *Betaproteobacteria* have been shown to be dominant in almost all DWDS studies published thus far. Many community composition studies have reported on variations in the dominance of these two classes depending on, but not limited to multiple factors, such as disinfection strategy (Gomez-Alvares *et al.,* 2012; Wang *et al*., 2014b) and seasonal variations (McCoy and VanBriesen, 2012 and 2014; Pinto *et al*., 2014; Prest *et al*., 2016b; Zlatanovic *et al*., 2017). DWDS bacterial communities are complex and vary across different DWDS as well as over time within a single DWDS (Proctor and Hammes, 2015). Some DWDSs show increased spatial and temporal variation whereas others have demonstrated microbial communities that remain spatially and temporally stable (Lautenschlager *et al*., 2013; Roeselers *et al.,* 2015).

In this study, while temporal trends the *Alpha-* and *Betaproteobacteria* were apparent, they were not significant. Although, seasonal variations between these proteobacterial classes have previously been reported (McCoy and VanBriesen, 2012 and 2014), it is important to note that the observed seasonal variations are also correlated to changes in disinfection residual concentrations and therefore should be considered in proper context when discussing observed temporal trends. Disinfectant dosing is often adjusted according to temperature shifts and therefore influences the microbial diversity (McCoy and VanBriesen, 2012). Temporal trends were more clearly reflected at an individual OTU level. The seasonal fluctuations in *Nitrosomonas*-like OTUs corresponded to potential increases in nitrification in the summer months, which correlated to increased temperatures and decreased chlorine residual levels.

While the bacterial community for the studied system was diverse, a small percentage of bacteria dominated the overall bacterial community. In this study, only 95 OTUs (~1% of all detected OTUs) had an overall relative abundance >1%, were shared across all samples over the course of the study and explained majority of the observed spatial-temporal variation. Gobet *et al*. (2010) demonstrated that ecological patterns were maintained after removal of 35-40% of rare sequences and beta diversity patterns were similar after denoising the data set, suggesting that the removal of rare sequences may be beneficial for data sets with a large

fraction of singletons. This was found to be the case in the current study as the subset of 95 OTUs showed temporal and spatial trends similar to those observed when all OTUs were considered, suggesting that the subset of OTUs may be sufficient to explain the variation in community structure and composition. This finding is not uncommon as Pinto *et al*. (2014) demonstrated that despite observing a wide taxonomic diversity, the changes in bacterial community structure could be explained by only 5% of the OTUs.

4.2 Effects of DWDS configuration and disinfection on bacterial community structure and composition

This DWDS consists of a complex network of underground pipelines with many interconnections, multiple reservoirs, and three different disinfection strategies (chlorination, chloramination and hypochlorination). The predominantly steel pipelines stretching over large distances (pipe length ranging from 6 km to 74 km), differ in internal lining material (bitumen or cement), diameter (ranging from 47 to 210 cm), and age (ranging from 16 to 81 years). The DWDS bulk water community therefore varies due to exposure to different disinfectants and the spatial heterogeneity of the system itself, allowing potential regrowth and increased interaction with biofilms present on the pipeline surfaces at a localised level (Srinivasan *et al*., 2008). In light of the variability within these parameters, the spatial heterogeneity of the DWDS can significantly influence the bacterial community as it moves through the DWDS. However, it is difficult to separate the effect of a single parameter on the bacterial community in a full-scale DWDS. For example, Wang *et al*. (2014a) reported on the combined effects of disinfectant, water age and pipe material creating different physicochemical conditions and ecological niches, which can then select and promote the growth of various microbes.

Despite the complexities of the system, we were able to show several key findings providing useful insights into the influence of the spatial structure of the DWDS on the bacterial community. Specifically, the spatial dynamics of the bacterial community conformed to the layout of the DWDS and the three disinfection strategies had a significant impact on the microbial community, particularly with the addition of the secondary disinfectant (chloramine). Together with increased retention time in a reservoir, chloramination caused a significant change in the bacterial community composition and structure. Although the impact of reservoirs on the microbial community is not fully understood, the observed change

in the microbial community was more pronounced in the first reservoir samples. These changes may be due to extended contact time between the microbial community and chloramine residuals due to potentially longer retention times in the reservoir (Hoefel *et al*., 2005). However, we were not able to decouple the effect of disinfection versus residence time in the reservoirs. This increased contact time with disinfectant residuals within the reservoirs coupled with residual decay, potentially allows for increased microbial regrowth during stagnation periods, causing shifts in abundances and in microbial community profiles (Lautenschlager *et al*., 2013). Furthermore, the observed change in community structure and membership correlated to the observed disinfectant decay within the chlorinated and chloraminated sections of the DWDS. However, more pronounced disinfectant decay was observed within the chloraminated section. Chloramine maintains extended disinfectant residuals, however within this section of the DWDS there are large distances between sample locations contributing to increased water age and disinfectant decay (Hoefel *et al*., 2005). In addition, the bacterial community displayed distance decay features with bacterial community structure becoming more dissimilar with an increase in distance between chloraminated sample locations. These findings may help in the modelling of the variability in the bacterial community as it moves away from the DWTP and through the DWDS (Schroeder *et al.,* 2015).

The effect of disinfection on the community composition has previously been reported, indicating that *Alphaproteobacteria* are typically dominant in both chlorinated and chloraminated water, whereas *Betaproteobacteria* were found to have increased abundance in chloraminated water as opposed to chlorinated water (Berry, 2006; Wang *et al*., 2014b). This was consistent with the observed changes in *Alpha-* and *Betaproteobacteria* in this study. A change in the bacterial community was also associated with an increase in richness following chloramination. Here, Baron *et al*., (2014) suggests that through chloramination, a potential loss of a select few dominant groups allowed for growth of a greater number of other species. Primary disinfection processes (chlorination in this case) typically dramatically reduce the bacterial community thereby potentially creating the opportunity of surviving microorganisms to proliferate and exploit the available nutrient pool (Hammes *et al*., 2008; Prest *et al*., 2016a).

4.3 Long-term seasonal variations in microbial community

In contrast to other studies (McCoy and VanBriesen, 2014; Pinto *et al*., 2014), winter months showed an increase in richness (observed OTUs), diversity and evenness although, an increase in richness in the winter months was also observed by Hwang and colleagues (2012). This increase in richness may be a consequence of a decrease in dominance of certain OTUs in the winter months, allowing for an increased detection of sequences of medium abundance and/or rare OTUs. Here a decrease in temperature may have a similar effect as chloramination in the reduction of dominant groups, allowing the detection and/or growth less dominant species. This decrease in dominance may also result in the observed increase in diversity and the OTUs being more evenly distributed. Gilbert *et al*., (2012) also observed this increase in richness in the winter months in marine microbial communities. Here the authors concluded that the observed seasonal changes in richness indicate that the most common and dominant bacterial taxa have temporally defined niches. Ling *et al*. (2016) also suggested that the observed seasonal variation influenced the dynamics of several core populations identified in DWDS biofilms and was main driver in the overall variation in the biofilm community.

Seasonal cycling was observed in the bacterial community structure and membership. However, the seasonal clustering observed by Pinto *et al*. (2014) was more well defined likely due to high variation in seasonal temperatures in Ann Arbor, Michigan USA and the change in the blend ratio of source waters (i.e. surface and ground water) in summer and winter. Significant seasonal variations were also observed by Prest *et al*. (2016b) in the Netherlands where seasonal temperatures vary by approximately 20 °C. Conversely, large temperature fluctuations rarely occur in South Africa and throughout the year and surface water is obtained from the same source. Despite this we did observe seasonal cycling, which could be a combination of temperature changes, changes in source water community composition, and changes in disinfectant residual concentrations over time. Temporal trends were, in fact, more clearly reflected within the bacterial community structure at individual sample points over the two years. Variability in community structure at each sample point over the duration of the study was considerable, specifically in the chloraminated section of the DWDS indicating a high level of temporal variation in bacterial community structure at a single sample location. This observed variability was largely due to the changes in abundance of dominant groups as similar trends were not clearly reflected in community membership.

Although small temporal trends were observed for community membership, the overall high level of dissimilarity in membership may be a consequence of the large number of rare OTUs, indicating that there is only a small percentage of shared membership at each location. In microbial ecology studies across a wide range of ecosystems, rare OTUs often contribute to a large proportion of the observed taxa. They represent a high diversity of bacterial and archaeal lineages and often explain the high levels of temporal variability in community membership (Shade *et al*., 2014; El-Chakhtoura *et al*., 2015). Although the ecological roles of the rare organisms are not well understood, it has been suggested that they may contribute to the community stability. In response to an environmental change, i.e. water temperature or disinfectant residuals, these rare OTUs may act as a reservoir and act as potential microbial seedbank when conditions change (Shade *et al*., 2014; El-Chakhtoura *et al*., 2015).

4.4 The interplay between spatial and temporal dynamics of the DWDS

A combination of parameters need to be considered for a comprehensive understanding of the bacterial community in the bulk water. When considering the DWDS as a whole, spatial dynamics explained more of the variation. This is likely due to the complexity of the large full-scale system such as this, with multiple interconnections, reservoirs and different disinfection residuals. It was shown that the extent of temporal variation in bacterial community structure, at each location, decreased as the bulk water moved away from the DWTP and through the DWDS, suggesting that the temporal dynamics are dampened by the spatial heterogeneity of the DWDS and the impact of DWDS specific microbial communities (i.e., sediments, biofilms) on the bulk water. This was clearly observed in the chloraminated section of the DWDS, where the distances between points were further apart than in the chlorinated and hypochlorinated sections.

Furthermore, samples farther away from treatment showed an increase in richness and diversity. Similar observations were reported by El-Chakhtoura and colleagues (2015), where the bacterial community structure changed during distribution, resulting in increased richness in the network. The increase of bacterial richness during distribution may be associated with regrowth, particularly in large distribution systems where bulk water is transported over long distances. Microbial growth in drinking water has been observed in the form of higher particle counts and increased turbidity (Liu *et al*., 2016), higher cell counts (Hammes *et al*.,

2008) and increase in the presence of indicator organisms in the final tap water compared to the water leaving the treatment plant (van der Wielen *et al*., 2016). The water leaving the treatment plant may therefore be impacted by the distribution system itself through processes such as pipe corrosion (Sun *et al.,* 2014), the detachment of biofilms (Chaves Simões and Simões, 2013) and suspension of loose deposits (Liu *et al*., 2013; Liu *et al*., 2017; Liu *et al*., 2018). These processes together in combination with increasing contact time with the disinfectant may explain the observed spatial dissimilarity observed in the bacterial community as bulk water moves away from the DWDP and through the DWDS.

At a more localised level, i.e., specific sample sites or sections, the temporal dynamics were clear and explained the variation as observed at individual sites/sections. This suggests that as the bulk water moves through the DWDS, the cumulative change in microbial community due to mechanisms ranging from growth, decay, sediment resuspension, and/or biofilm detachment is larger than the temporal change over a sampling frequency of this study (i.e., monthly). It is plausible that biofilms and/or sediments in the studied DWDS exhibit much higher temporal stability as compared to the bulk water and their seeding of the bulk water is the primary mechanism for the decreasing temporal variability with increasing DWDS distance from the DWTP.

5. Conclusions

Through conducting a long-term survey spanning two years, we were able to comprehensively characterise the temporal and spatial dynamics of the microbial community within a complex, large-scale DWDS. Here we show strong temporal trends in richness and diversity, correlated with seasonal changes in disinfectant residuals as well as seasonal cycling in the bacterial community structure and composition in the DWDS. Temporal trends were dominant at a localised level, showing seasonal variations, but when considering the DWDS in its entirety, spatial dynamics were stronger and explained more of the variation in the bacterial community structure. This study highlighted the interplay between the spatial and temporal dynamics of the DWDS. Here, temporal dynamics decreased as bulk water moved away from the treatment plant due to the potential seeding of the bulk water by the relatively temporally stable communities (i.e. biofilms and loose deposits) inherent to the DWDS. Complete understanding of the factors driving the changes in large-scale DWDS bacterial communities may be difficult to achieve as these DWDS are complex and inherently

dynamic. However, through a long-term, high-frequency investigation such as this, we were able to clearly observe seasonal and annual patterns in the microbial community that would have otherwise been missed in a short-term study. Even though high diversity and variation was observed within the microbial community, we detected a core community that was present in all samples collected as part of this study. This core community was able to tolerate a range of physical-chemical variations within the system, Therefore, this study contributes to current knowledge base in this field and provides the opportunity for drinking water utilities to understand the range of mechanisms that influence the bacterial community structure and composition, over varying temporal scales and/or operational stages.

6. References

1. Anderson, M.J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*. 26 (1): 32-46.

2. Baron, J. L., Vikram, A., Duda, S., Stout, J. E. and Bibby, K. (2014). Shift in the microbial ecology of a hospital hot water system following the introduction of an on-site monochloramine disinfection system. *PLOS One.* 9(7): e102679.

3. 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.

4. Berry, D., Xi, C. and Raskin, L. (2006). Microbial ecology of drinking water distribution systems. *Current Opinion Biotechnology.* 17: 297-302.

5. Chaves Simões, L. and Simões, M. (2013). Biofilms in drinking water: problems and solutions. *RSC Advances*. 3(8): 2520-2533.

6. Dinno, A. (2017). dunn.test: Dunn's Test of multiple comparisons using rank sums. R package version 1.3.3. (http://CRAN.R-project.org/package=dunn.test).

7. DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P. and Andersen, G. L. (2006). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology*. 72: 5069-72.

8. Douterelo, I., Husband, S. and Boxall, J. B. (2014). The bacterial composition of biomass recovered by flushing an operational drinking water distribution system. *Water Research*. 54: 100-114.

9. Edgar, R.C., Haas, B. J., Clemente, J. C., Quince, C. and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*. doi: 10.1093/bioinformatics/btr381.

10. El-Chakhtoura, J., Prest, E., Saikaly, P., van Loosdercht., Hammes, F. and Vrouwenvedler, H. (2015). Dynamics of bacterial communities before and after distribtuion in a full-scale drinking water network. *Water Research*. 74: 180-190.

11. Evans, J., Sheneman, L. and Foster, J. A. (2006). Relaxed Neighbour-Joining: A Fast Distance-Based Phylogenetic Tree Construction Method. *Journal of Molecular Evolution*. 62: 785-792.

12. Excoffier, L. (1993). Analysis of molecular variance (AMOVA) version 1.55. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

13. Fox, J. and Weisberg, S. (2011). An {R} Companion to Applied Regression, Second Edition. Thousand Oaks CA: Sage.
URL:http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

14. Gall, A. M., Mariñas, B. J., Lu, Y. and Shisler, J. L. (2015). Waterborne viruses: A barrier to safe drinking water. *PLoS Pathogens*. 11(6): e1004867. doi: 10.1371/journal.ppat.1004867.

15. Gilbert, J. A., Steele, J. A., Caporaso, J. G., Steinbrück., Reeder, J., Temperton, B., Huse, S., McHardy, A. C., Knight, R., Joint, I., Somerfield, P., Fuhrman, J. A. and Field, D. (2012). Defining seasonal marine microbial community dynamics. *The ISME Journal*. 6: 298-308.

16. Gobet, A., Quince, C. and Ramette, A. (2010). Multivariate cuttoff level analysis (MultiCoLA) of large community data sets. *Nucleic Acids Research*. 38(15): e155. doi:10.1093/nar/gkq545.

17. Gomez-Alvarez, V., Revetta, R. P. and Santo Domingo, J. W. (2012). Metagenomic analysis of drinking water receiving different disinfection treatments. *Applied and Environmental Microbiology*. 78(17): 6095-6102.

18. 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.

19. Henne, K., Kahisch, L., Brettar, I. and Holfe, M. G. (2012). Analysis of structure and composition of bacterial core communities in mature drinking water biofilms and bulk water of a citywide network in Germany. *Applied and Environmental Microbiology*. 78(10): 3530-3538.

20. Hoefel, D., Monis, P. T., Grooby, W. L., Andrews, S. and Saint, C .P. (2005). Culture-independent techniques for repid detection of bacteria associated with the loss of chloramine residual in a drinking water system. *Applied and Environmental Microbiology*. 71(11): 6479-6488.

21. 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.

22. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. and 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: 5112-5120.

23. 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.

24. Lautenschlager, K., Hwang, C., Ling, F., Liu, 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.

25. Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J. and Knight, R. (2011). UniFrac: an effective distance metric for microbial community comparison. *The ISME journal*. 5(2): 169-172.

26. Ling, F., Hwang, C., LeChevallier, M. W., Anderson, G. L. and Liu, W-T. (2016). Core-satellite populations and seasonality of water meter biofilms in a metropolitan drinking water distribution system. *The ISME Journal*. 10: 582-595.

27. 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.

28. Liu, G., Bakker, G. L., Vreeberg, J. H. G., Verberk, J. Q. J. C., Medama, G. J., Liu, M. C. and Van Dijk, J. C. (2014). Pyrosequencing reveals bacterial communities in unchlorinated drinking water distribution system: An integral study of bulk water, suspended solids, loose deposits and pipe wall biofilm. *Environmental Science and Technology*. 48: 5467-5476.

29. Liu, G., Ling, F. Q., Van der Mark, E. J., Zhang, X. D., Knezev, A., Verberk, J. Q. J. C., Van der Meer, W. G. J., Medema, G. J., Liu, W. T. and Van Dijk, J. C. (2016). Comparison of particle-associated bacteria from a drinking water treatment plant and distribution reservoirs with different water sources. *Scientific Reports*. 6: 20367.

30. Liu, G., Zhang, Y., Knibbe, W-J., Feng, C., Liu, W., Medema, G. and van der Meer, W. (2017). Potential impacts of changing water suppy-water quality on drinking water distribution: A review. *Water Research*. 116: 135-148.

31. Liu, G., Zhang, Y., Van der Mark, E., Magic-Knezev, A., Pinto, A., van den Bogert, B., Liu, W., Van der Meer, W. and Medema, G. (2018). Assessing the origin of bacteria in tap water and distribution system in an unchlorinated drinking water system by SourceTracker using microbial community fingerprints. *Water Research*. 138: 86-96.

32. McCoy, S. T. and VanBriesen, J. M. (2012). Temporal variability of bacterial diversity in a chlorinated drinking water distribution system. *Journal of Environmental Engineering*. 138(7): 786-795.

33. Mccoy, S. T. and Van Briesen, J. M. (2014). Comparing spatial and temporal diversity of bacteria in a chlorinated drinking water distribution system. *Environmental Engineering Science*. 31(1): 32-41.

34. 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.

35. Nescerecka, A., Rubulis, J., Vital, A., Juhna, T. and Hammes, F. (2014). Biological instability in a chlorinated drinking water distribution system. *PLOS One*. 9(5): e96354.

36. Niquette, P., Servais, P. and Savoir, R. (2000). Impacts of pipe materials on densities of fixed bacterial biomass in a drinking water distribution system. *Water Research*. 34(6):1952-1956.

37. Oksanen, J., Guillaume Blanchet, F., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, H. H. and Wagner, H. (2015). Vegan: Community Ecology Package. R package version 2.3-0. (http://CRAN.R-project.org/package=vegan).

38. 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.

39. 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 ASM*. 5(3): e01135-14.

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

41. Prest, E. I., Weissbrodt, D. G., Hammes, F., van Loosdrecht, M. C. M. and Vrouwenvelder, J. S. (2016b). Long-term bacterial dynamics in a full scale drinking water distribution system. *PLoS ONE*. DOI: 10.1371/journal.pone.0164445.

42. Proctor, C. R. and Hammes, F. (2015). Drinking water microbiology – from measurement to management. *Current Opinion in Microbiology*. 33: 87-95.

43. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*. 41(D1): D590-D596.

44. R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/.

45. 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.

46. Schrader, C., Schielke, A., Ellerbroek, L. and Johne, R. (2012). PCR inhibitors – occurrence, properties and removal. *Journal of Applied Microbiology*. 113: 1014-1026.

47. 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. and 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): 7537-7541.

48. Schroeder, J. L., Lunn, M., Pinto, A. J., Raskin, L. and Sloan, W. T. (2015). Probabilistic models to describe tge dynamics of migrating microbial communities. *PLoS ONE*. DOI:10.1371/journal.pone.0117221.

49. Shade, A., Jones, S. E., Caporaso, J. G., Handelsman, J., Knight, R., Fierer, N. and Gilbert, J. A. (2014). Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *mBio ASM*. 5(4): 1-9.

50. Shaw, J. L. A., Monis, P., Weyrich, L. S., Sawade, E., Drikas, M. and Cooper, A. J. (2015). Using amplicon sequencing to characterize and monitor bacterial diversity in drinking water distribution systems. *Applied and Environmental Microbiology*. 81(18): 6463-6473.

51. Siqueira, V. M. and Lima, N. (2013). Biofilm formation by filamentous fungi recovered from a water system. *Journal of Mycology*. 1-9.

52. Srinivasan, S., Harrington, G. W., Xagoraraki, I. and Goel, R. (2008). Factors affecting bulk to total bacteria ratio in drinking water distribution systems. *Water Research*. 42: 3393-33404.

53. Sun, H., Shi, B., Lytle, D. A., Bai, Y. and Wang, D. (2014). Formation and release behavior of iron corrosion products under the influence of bacterial communities in a simulated water distribution system. *Environmental Science: Processes and Impacts*. 16(3): 576-585.

54. Thomas, J. M. and Ashbolt, N. J. (2011). Do Free-Living amoebae in treated drinking water systems present an emerging health risk? *Environmental Science and Technology*. 45(3): 860-869.

55. Urakawa, H., Martens-Habbena, W. and Stahl, D. A. (2010). High abundance of ammonia-oxidising Archaea in coastal waters, determined by using a modified DNA extraction method. *Applied and Environmental Microbiology*. 76(7): 2129-2135.

56. Van Aken, B. and Lin, L-S. (2011). Effect of the disinfection agents chlorine, UV irradiation, silver ions, and TiO_2 nanoparticles/near-UV on DNA molecules. *Water Science and Technology*. 64(6): 1226-1232.

57. van der Wielen, P. W. J. J., Bakker, G., Atsma, A., Lut, M., Roeselers, G. and Graaf, B. D. (2016). A survey of indicator parameters to monitor regrowth in unchlorinated drinking water. *Environmental Science: Water Research and Technology*. 4(2): 683-692.

58. Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, http://ggplot2.org.

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

60. Wang, H., Proctor, C. R., Edwards, M. A., Pryor, M., Santo Domingo, J. W., Ryu, H., Camper, A. K., Olson, A. and Pruden, A. (2014b). Microbial community response to chlorine conversion in a chloraminated drinking water distribution system. *Environmental Science and Technology*. 48: 10624-10633.

61. Wu, H., Zhang, J., Mi, Z., Xie, S., Chen, C. and Zhang, X. (2015). Biofilm bacterial communities in urban drinking water distribution systems transporting waters with different purification strategies. *Environmental Biotechnology*. 99: 1947-1955.

62. You, J., Das, A., Dolan, E. M. and Hu, Z. (2009). Ammonia-oxidising archaea involved in nitrogen removal. *Water Research*. 43: 1801-1809.

Zlatanovic, L., van der Hoek, J. P. and Vreeburg, J. H. G. (2017). An experimental study on the influence of water stagnation and temperature change on water quality in a full-scale domestic drinking water system. *Water Research*. 123: 761-772.

Tables

Table 1: Outline and description of the multiple DWDS sample locations included in the study over the two-year study period and the number of samples sequenced for each sample site

	DWDS sample location	Sample site*	Number of samples sequenced	Sample location description
1	CHL1	CHL1	15	Chlorinated water leaving the treatment plant
2	CHL2	CHL2	13	Chlorinated water entering boosting station
3	CHM1	CHM1.1	17	Chloraminated water leaving the boosting station
4		CHM1.2	14	Chloraminated water leaving the boosting station
5	CHM2	CHM2.1	22	Chloraminated reservoir
6		CHM2.2	20	Chloraminated reservoir
7	CHM3	CHM3.1	21	Chloraminated water leaving the pumping station
8		CHM3.2	20	Chloraminated water leaving the pumping station
9		CHM3.3	21	Chloraminated water leaving the pumping station
10	CHM4	CHM4.1	19	Chloraminated water pipeline
11		CHM4.2	21	Chloraminated water pipeline
12		CHM4.3	20	Chloraminated water pipeline
13	CHM5	CHM5.1	21	Chloraminated reservoir
14		CHM5.2	22	Chloraminated reservoir
15		CHM5.3	23	Chloraminated reservoir

16	HCHL1	HCHL1	21	Hypochlorinated water leaving boosting station
17	HCHL2	HCHL2	21	Hypochlorinated water leaving boosting station
18	HCHL3	HCHL3	17	Hypochlorinated water pipeline

* Some locations included multiple sample sites

Table 2: Percentage of total sequences, mean relative abundance and standard deviation of the most abundant OTUs (percentage of total sequences > 1%)

Operational taxonomic unit	Taxonomic classification	Percentage of total sequences	Percentage MRA*	SD*
OTU1	Genus: *Nitrosomonas*	18.49	14.97	18.94
OTU2	Order: *Rhizobiales*	14.7	14.45	19.92
OTU3	Order: *Rhizobiales*	12.47	12.11	14.07
OTU4	Genus: *Shingomonas*	9.72	9.16	10.02
OTU5	Family: *Gemmataceae*	3.72	4.81	7.54
OTU6	Genus: *Hyphomicobium*	1.98	1.8	3.34
OTU7	Genus: *Nitrospira*	1.83	1.68	3.74
OTU8	Class: *Betaproteobacteria* - SBla14	1.77	1.63	4.61
OTU9	Genus: *Planctomyces*	1.23	0.97	4.3

* MRA: Mean relative abundance

* SD: Standard deviation

Figures
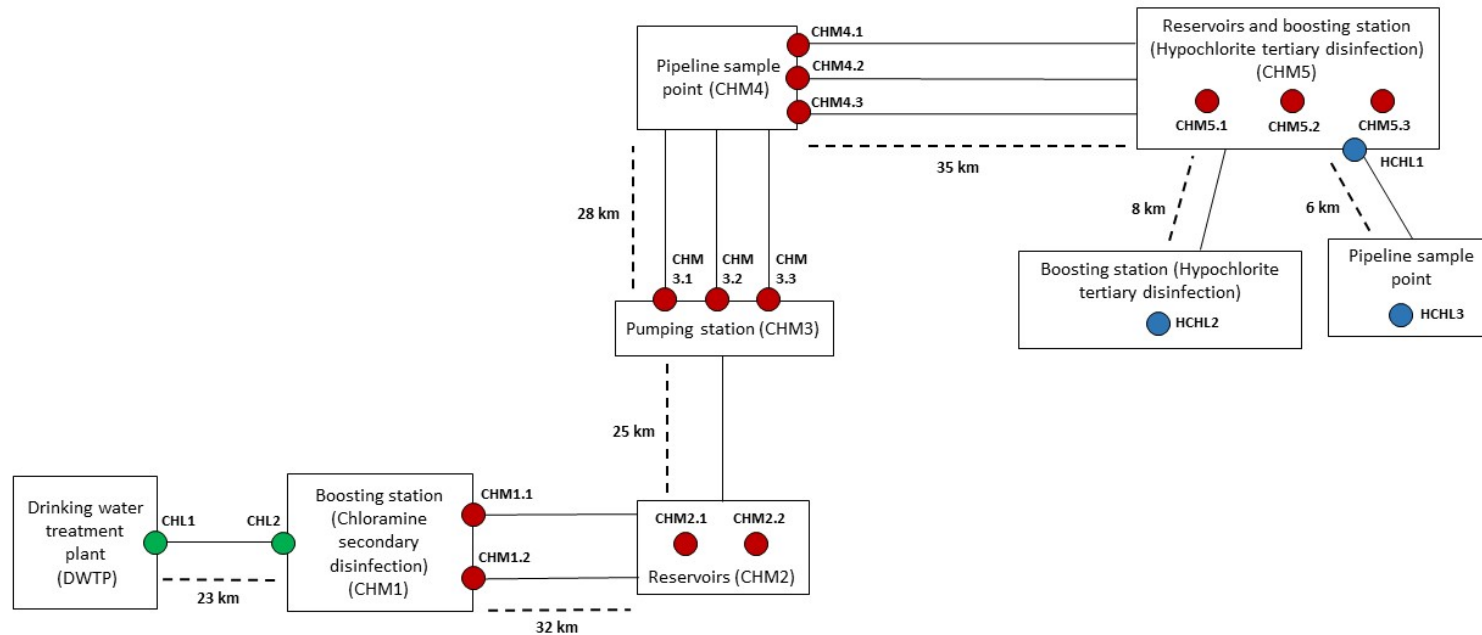
Figure 1: Schematic showing the layout of DWDS sample locations and respective sampling sites included in this study. Sample sites are indicated as circles and coloured according to the disinfectant residual used in that section of the DWDS [green circles, chlorine (CHL); red circles, chloramine (CHM) and blue circles, hypochlorite (HCHL)]. Dashed lines indicate the pipe distances (km) between sample locations.
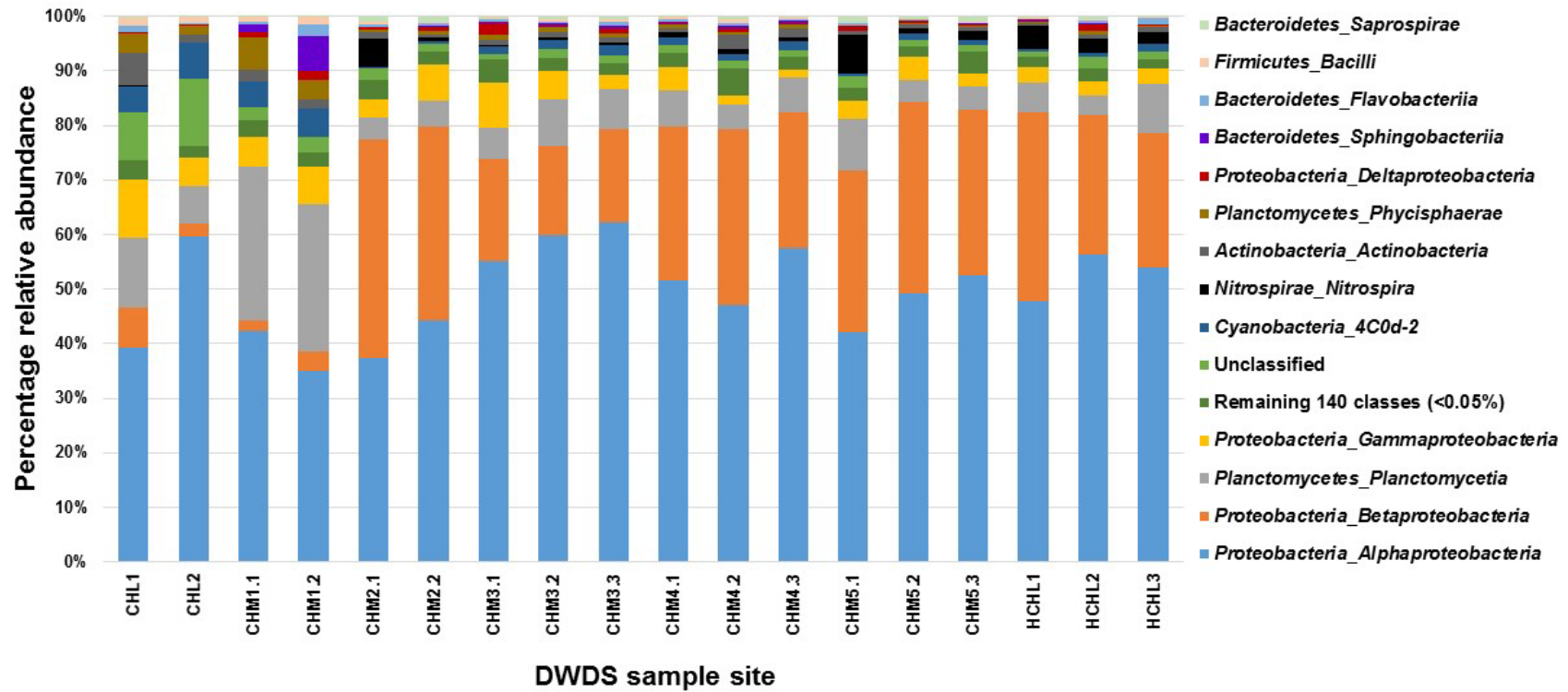
Figure 2: Class-level relative abundances of all bacterial sequences detected across the duration of the study at each DWDS sample site. The top thirteen most abundant bacterial classes are shown with the remaining 140 (constituting < 0.05% of the total abundance) grouped as a single group. Classes together with the phylum they belong to (phylum_class) are shown in the legend on the right.
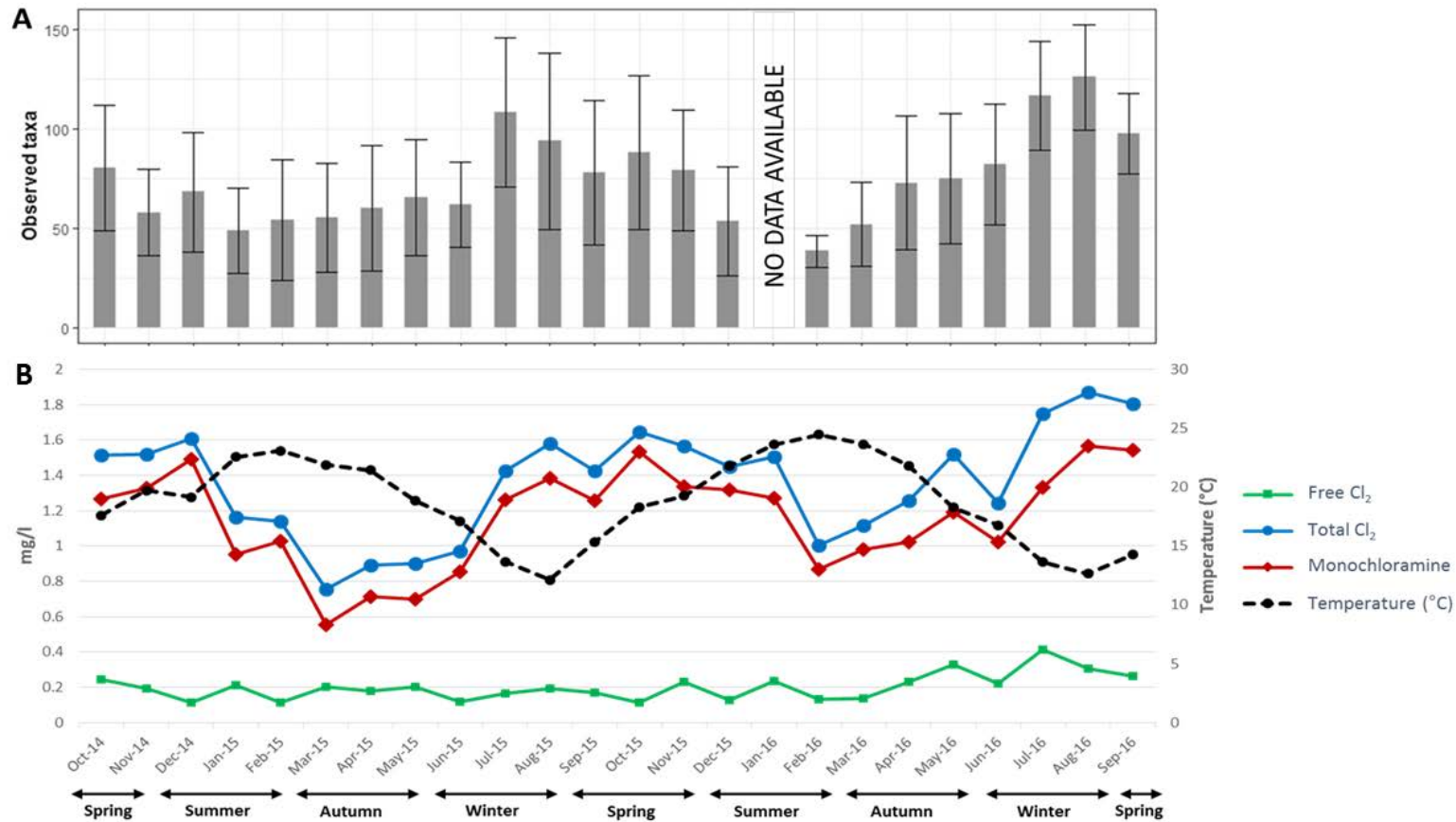
Figure 3: Temporal change in richness (observed taxa) averaged across all sample locations for which chemical data were available (i.e. CHL2, CHM1 – 4, CHM5.3 and HCHL1 – 3) for each month (A), correlated with average temperature (dashed black line) and average concentrations of disinfectant residuals (i.e., free chlorine [green squares], total chlorine [blue circles] and monochloramine [red diamonds]) over the duration of the study (B). Error bars represent the standard deviation in observed taxa across all sample sites within each month.
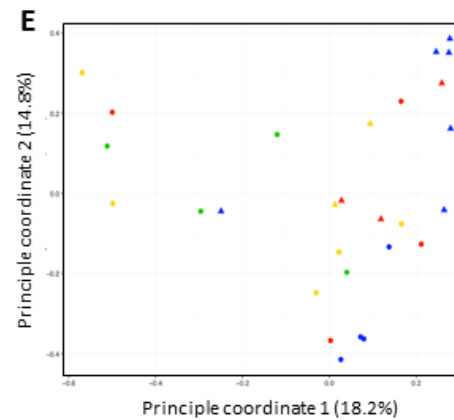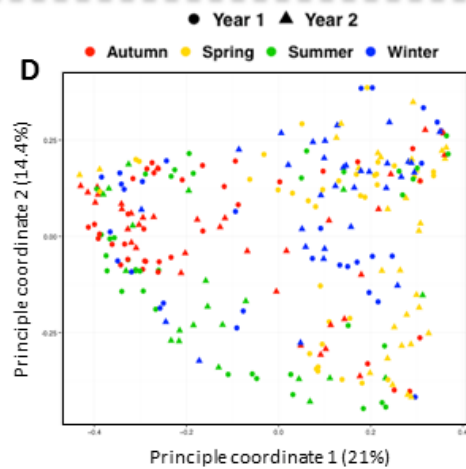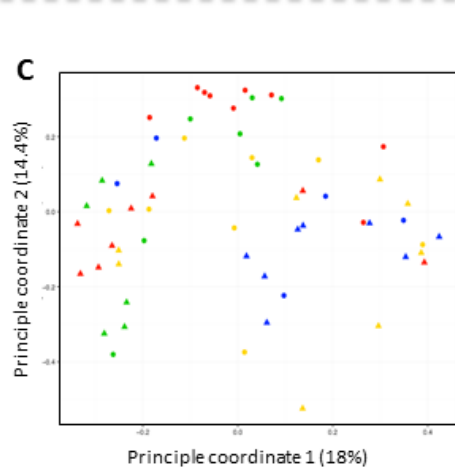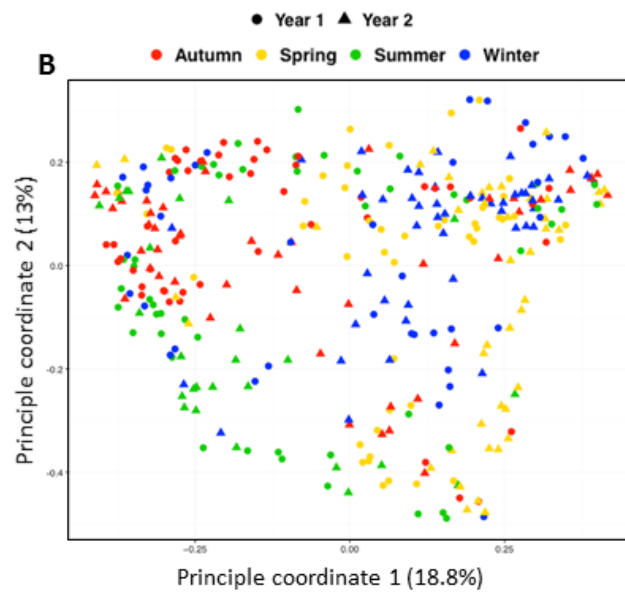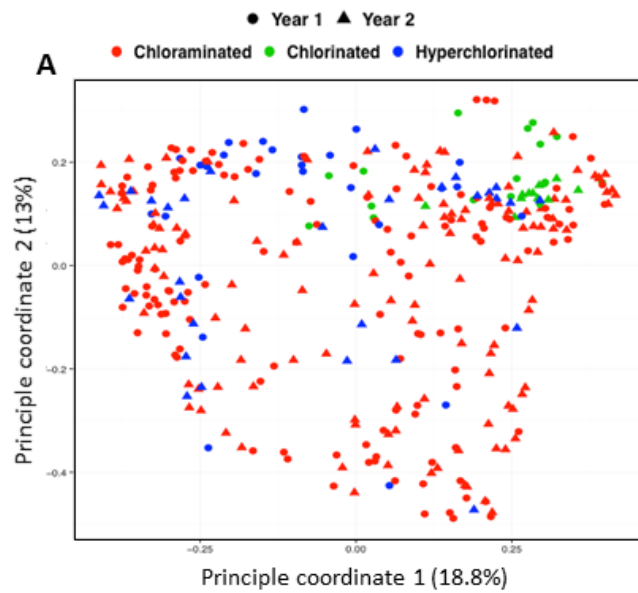
Figure 4: Principal-coordinate analyses plots showing the spatial and temporal variability of the bacterial community structure within the DWDS using Bray-Curtis dissimilarities. Spatial groupings are shown in plot (A) where data points are coloured based on disinfection strategy and shaped based on year. Temporal groupings are shown in plot (B) where data points are coloured based on season and shaped based on year. Colour and shapes are indicated in the legends on the top of both plots. The three lower plots indicate the temporal groupings within the three different disinfection sections, i.e. chlorinated (C), chloraminated (D) and hypochlorite (E). These three plots are coloured by season and shaped based on year, shown in the legend on the top of the plot D.
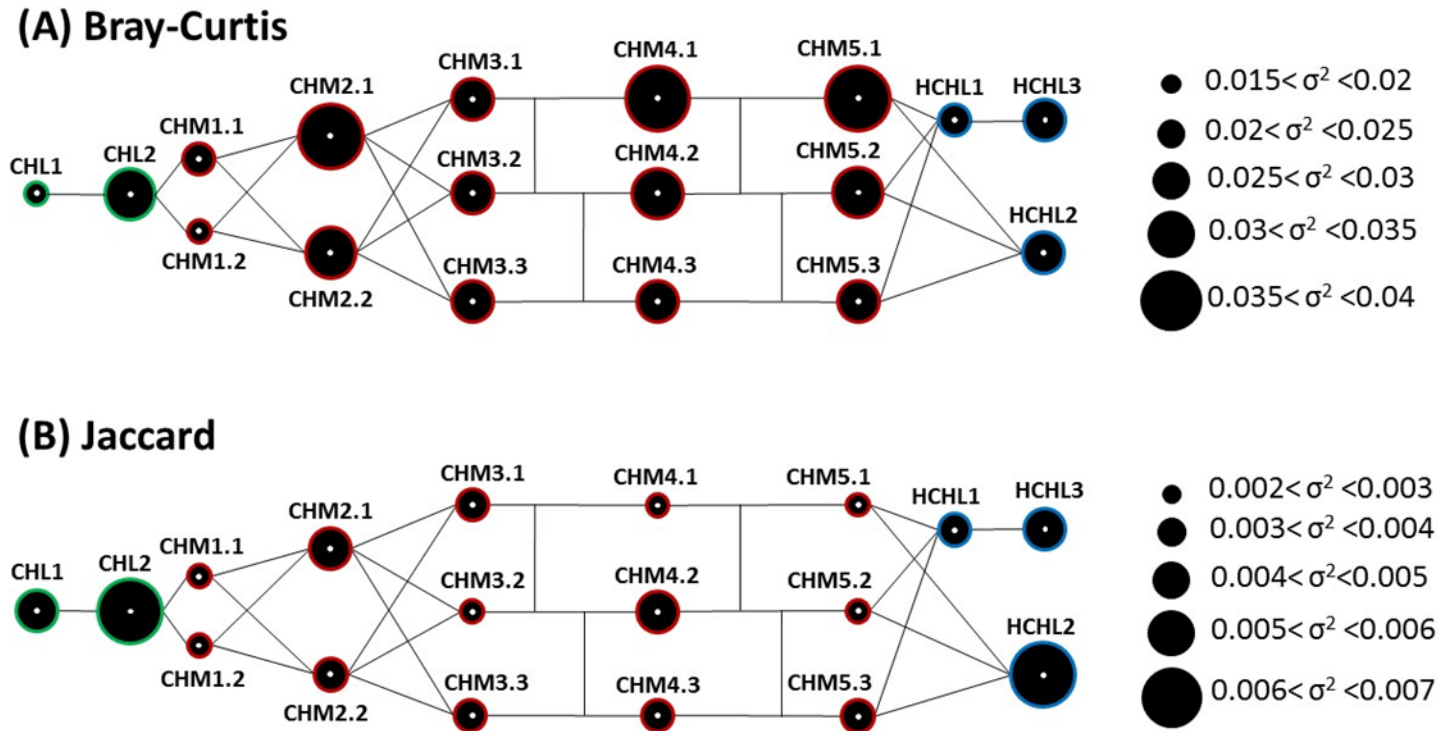
Figure 5: A schematic following the layout of the DWDS showing the extent of temporal variation at each sample site. White dots at the centre of each circle represent the median of all pairwise Bray-Curtis dissimilarity distances (A) and Jaccard distances (B) within each sample site over the duration of the study. Sizes of the black circles indicate the extent of temporal variance at each sample site. The extent of temporal variance is indicated in the legend on the right of the figure. Sample sites are coloured according to disinfection strategy (chlorination [green], chloramination [red] and hypochlorination [blue]).
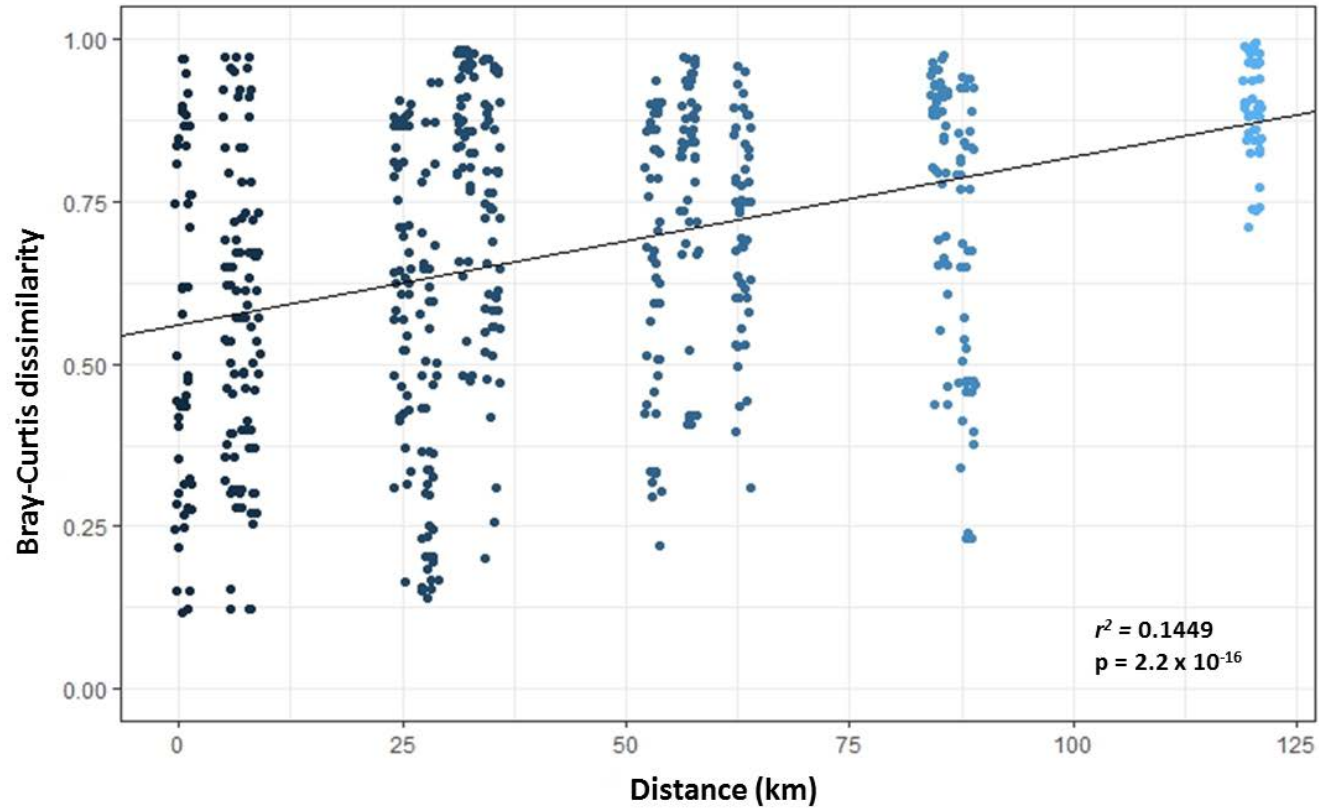
Figure 6: Distance decay features of the chloraminated section of the DWDS (CHM1 – CHM5) using pair wise structure-based Bray-Curtis distances. The line through the graph indicates the linear regression model. r^2 and significance values are shown on the graph.
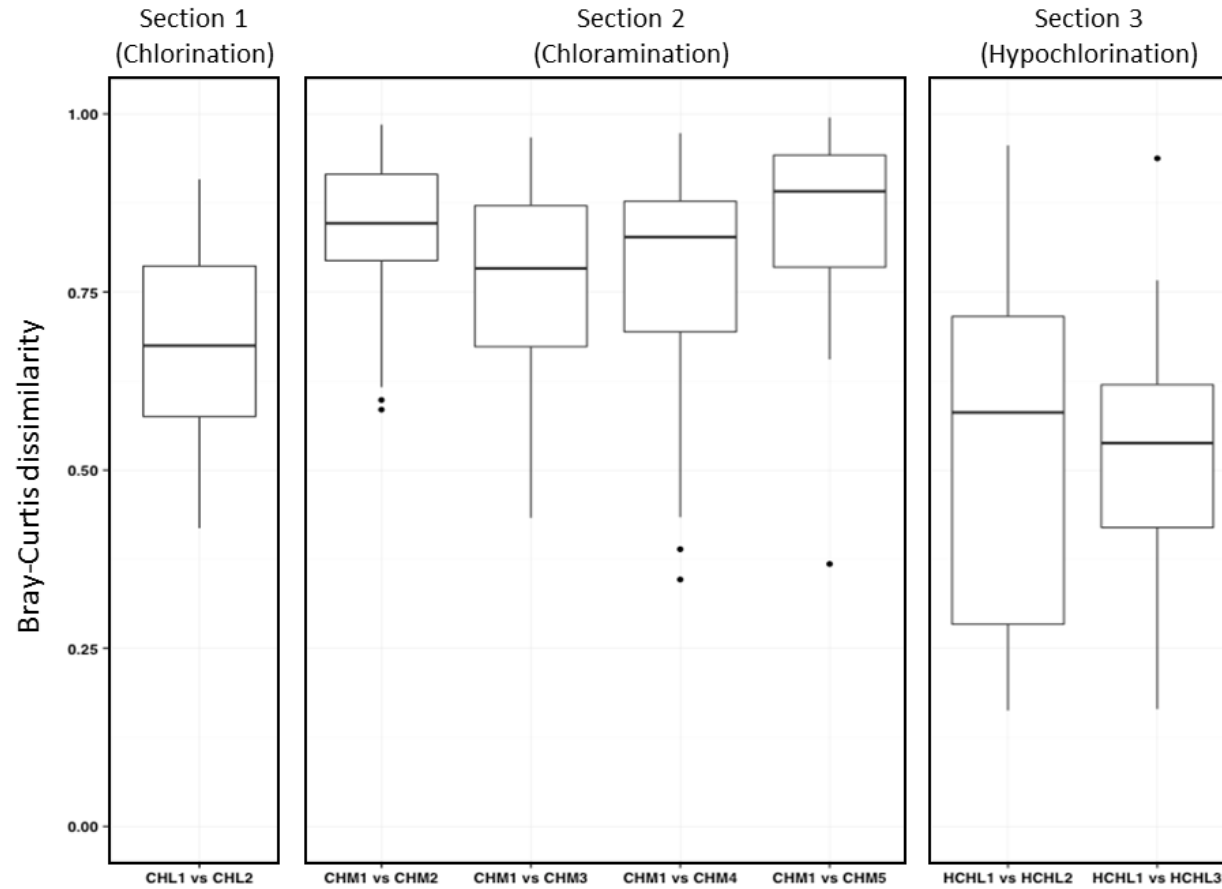
Figure 7: Comparisons of beta diversity distances between samples immediately after disinfection and with other samples within each disinfection section were performed for all beta-diversity metrics. Pairwise distances were included from samples within a location from the same month.

Appendix A: Chapter 2 supplementary information

Supplementary tables

Table A1: Samples excluded due to failed sequencing

Sample site	Month	Season
CHL1	Nov-14	Spring
	Jan-15	Summer
	Feb-15	Summer
	Aug-15	Winter
	Dec-15	Summer
	Feb-16	Summer
	Mar-16	Autumn
	Sep-16	Spring
CHL2	Apr-15	Autumn
	May-15	Winter
	Jul-15	Winter
	Oct-15	Spring
	Nov-15	Spring
	Dec-15	Summer
	Feb-16	Summer
	Mar-16	Autumn
	May-16	Winter
	Sep-16	Spring
CHM1.1	Feb-15	Summer
	Jul-15	Winter
	Dec-15	Summer
	Feb-16	Summer
	Mar-16	Autumn
	Aug-16	Winter
CHM1.2	Apr-15	Autumn
	May-15	Winter
	Jul-15	Winter
	Sep-15	Spring

	Oct-15	Spring
	Nov-15	Spring
	Dec-15	Summer
	Feb-16	Summer
	Aug-16	Winter
CHM2.1	Apr-15	Autumn
CHM2.2	Jul-15	Winter
	Jul-16	Winter
	Sep-16	Spring
CHM3.1	Jun-15	Winter
	Jul-15	Winter
CHM3.2	Nov-15	Spring
	Dec-15	Summer
	Jul-16	Winter
CHM3.3	Aug-15	Winter
	Sep-15	Spring
CHM4.1	Aug-15	Winter
	Sep-15	Spring
	Apr-16	Autumn
	Jun-16	Winter
CHM4.2	Aug-15	Winter
	Sep-15	Spring
CHM4.3	Aug-15	Winter
	Sep-15	Spring
	Sep-16	Spring
CHM5.1	Jun-16	Winter
	Sep-16	Spring
CHM5.2	Aug-15	Winter
CHM5.3	-	-
HCHL1	Jul-15	Winter
	Aug-15	Winter
HCHL2	Feb-15	Summer
	Sep-16	Spring
HCHL3	Nov-14	Spring
	Feb-15	Summer

	Jun-15	Winter
	Jul-15	Winter
	Apr-16	Autumn
	Jun-16	Winter

Table A2: Taxonomy of the subset of OTUs (95 OTUs) identified as the core community

OTU	Number of sequences per OTU	Phylum	Class	Order	Family	Genus	Species
Otu00001	1537994	*Proteobacteria*	*Betaproteobacteria*	*Nitrosomonadales*	*Nitrosomonadaceae*	*Nitrosomonas*	*Nitrosomonas oligotropha*
Otu00002	1223010	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	unclassified	unclassified	unclassified
Otu00003	1037491	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	unclassified	unclassified	unclassified
Otu00004	809246	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Sphingomonas*	*Sphingomonas asaccharolytica*
Otu00005	310577	*Planctomycetes*	*Planctomycetia*	*Gemmatales*	*Gemmataceae*	unclassified	unclassified
Otu00006	164355	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Hyphomicrobiaceae*	*Hyphomicrobium*	unclassified
Otu00007	152399	*Nitrospirae*	*Nitrospira*	*Nitrospirales*	*Nitrospiraceae*	*Nitrospira*	unclassified
Otu00008	147065	*Proteobacteria*	*Betaproteobacteria*	SBla14	unclassified	unclassified	unclassified
Otu00009	102091	*Planctomycetes*	*Planctomycetia*	*Planctomycetales*	*Planctomycetaceae*	*Planctomyces*	unclassified
Otu00010	74088	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Hyphomicrobiaceae*	*Hyphomicrobium*	unclassified
Otu00011	71958	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Hyphomicrobiaceae*	*Hyphomicrobium*	unclassified
Otu00012	71515	*Proteobacteria*	*Betaproteobacteria*	*Methylophilales*	*Methylophilaceae*	*Methylotenera*	*Methylotenera mobilis*
Otu00013	63553	*Proteobacteria*	*Betaproteobacteria*	*Gallionellales*	*Gallionellaceae*	*Gallionella*	unclassified
Otu00014	60813	*Proteobacteria*	*Betaproteobacteria*	unclassified	unclassified	unclassified	unclassified
Otu00015	58212	*Proteobacteria*	*Betaproteobacteria*	unclassified	unclassified	unclassified	unclassified
Otu00017	53380	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Sphingobium*	*Sphingobium yanoikuyae*
Otu00018	51323	*Proteobacteria*	*Alphaproteobacteria*	*Rhodospirillales*	*Acetobacteraceae*	unclassified	unclassified
Otu00019	48399	*Cyanobacteria*	4C0d-2	MLE1-12	unclassified	unclassified	unclassified
Otu00020	47207	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	unclassified	unclassified
Otu00021	46987	*Proteobacteria*	*Betaproteobacteria*	unclassified	unclassified	unclassified	unclassified
Otu00022	46057	*Planctomycetes*	*Planctomycetia*	*Planctomycetales*	*Planctomycetaceae*	*Planctomyces*	unclassified

Otu00023	42443	*Proteobacteria*	*Betaproteobacteria*	unclassified	unclassified	unclassified	unclassified
Otu00024	42069	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Erythrobacteraceae*	*Erythromicrobium*	*Erythromicrobium ramosum*
Otu00025	42021	*Proteobacteria*	*Alphaproteobacteria*	unclassified	unclassified	unclassified	unclassified
Otu00026	38929	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*Rhodobacter*	unclassified
Otu00027	38460	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Hyphomonadaceae*	unclassified	unclassified
Otu00028	38080	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	ACK-M1	unclassified	unclassified
Otu00029	37000	*Proteobacteria*	*Alphaproteobacteria*	unclassified	unclassified	unclassified	unclassified
Otu00030	35315	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*Acidovorax*	unclassified
Otu00031	33238	*Proteobacteria*	*Alphaproteobacteria*	*Caulobacterales*	*Caulobacteraceae*	*Mycoplana*	unclassified
Otu00032	32369	*Planctomycetes*	*Planctomycetia*	*Pirellulales*	*Pirellulaceae*	unclassified	unclassified
Otu00033	32048	*Planctomycetes*	*Phycisphaerae*	*Phycisphaerales*	unclassified	unclassified	unclassified
Otu00034	31501	unclassified	unclassified	unclassified	unclassified	unclassified	unclassified
Otu00035	30663	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	unclassified	unclassified
Otu00037	29158	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Bradyrhizobiaceae*	unclassified	unclassified
Otu00038	28463	*Planctomycetes*	*Phycisphaerae*	*Phycisphaerales*	unclassified	unclassified	unclassified
Otu00039	27107	*Planctomycetes*	*Planctomycetia*	*Gemmatales*	*Isosphaeraceae*	unclassified	unclassified
Otu00040	26689	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*	unclassified
Otu00041	25564	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacteriales*	*Enterobacteriaceae*	*Escherichia*	*Escherichia coli*
Otu00042	25093	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	unclassified	unclassified	unclassified
Otu00043	24912	*Bacteroidetes*	*[Saprospirae]*	*[Saprospirales]*	*Chitinophagaceae*	*Sediminibacterium*	unclassified
Otu00044	24849	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	unclassified	unclassified
Otu00046	23570	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Hyphomicrobiaceae*	*Hyphomicrobium*	unclassified
Otu00047	23033	*Proteobacteria*	*Alphaproteobacteria*	*Rhodospirillales*	*Rhodospirillaceae*	*Reyranella*	*Reyranella massiliensis*
Otu00048	22471	*Proteobacteria*	*Betaproteobacteria*	*Gallionellales*	*Gallionellaceae*	*Gallionella*	unclassified
Otu00049	22436	*Proteobacteria*	*Gammaproteobacteria*	unclassified	unclassified	unclassified	unclassified
Otu00050	21022	*Proteobacteria*	*Gammaproteobacteria*	*Alteromonadales*	*[Chromatiaceae]*	*Rheinheimera*	unclassified

Otu00051	19869	*Acidobacteria*	*Holophagae*	*Holophagales*	*Holophagaceae*	unclassified	unclassified
Otu00052	19310	*Cyanobacteria*	4C0d-2	MLE1-12	unclassified	unclassified	unclassified
Otu00053	17669	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	unclassified	unclassified	unclassified
Otu00054	16870	*Proteobacteria*	*Betaproteobacteria*	*Rhodocyclales*	*Rhodocyclaceae*	*Sulfuritalea*	unclassified
Otu00055	16701	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Novosphingobium*	*Novosphingobium stygium*
Otu00056	16690	*Cyanobacteria*	4C0d-2	MLE1-12	unclassified	unclassified	unclassified
Otu00057	16405	*Crenarchaeota*	*Thaumarchaeota*	*Cenarchaeales*	*Cenarchaeaceae*	unclassified	unclassified
Otu00058	15938	*Proteobacteria*	*Alphaproteobacteria*	unclassified	unclassified	unclassified	unclassified
Otu00059	15266	*Proteobacteria*	*Betaproteobacteria*	*Rhodocyclales*	*Rhodocyclaceae*	*Rhodocyclus*	unclassified
Otu00060	15064	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	unclassified	unclassified
Otu00061	14218	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*Rhodobacter*	unclassified
Otu00062	14203	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Methylobacteriaceae*	*Magnetospirillum*	*Magnetospirillum magnetotacticum*
Otu00064	13980	*Proteobacteria*	*Deltaproteobacteria*	*Spirobacillales*	unclassified	unclassified	unclassified
Otu00065	12429	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	ACK-M1	unclassified	unclassified
Otu00066	12288	*Bacteroidetes*	*Flavobacteriia*	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*	*Flavobacterium succinicans*
Otu00067	12288	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*	unclassified
Otu00068	12254	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Mycobacteriaceae*	*Mycobacterium*	unclassified
Otu00069	11817	*Proteobacteria*	*Gammaproteobacteria*	*Aeromonadales*	*Aeromonadaceae*	*Aeromonas*	unclassified
Otu00070	11755	*Proteobacteria*	*Alphaproteobacteria*	*Rhodospirillales*	*Acetobacteraceae*	unclassified	unclassified
Otu00071	11689	*Bacteroidetes*	*Sphingobacteriia*	*Sphingobacteriales*	*Sphingobacteriaceae*	*Pedobacter*	unclassified
Otu00072	11456	*Proteobacteria*	*Gammaproteobacteria*	*Alteromonadales*	*Alteromonadaceae*	*Cellvibrio*	unclassified
Otu00073	11122	*Proteobacteria*	*Betaproteobacteria*	*Nitrosomonadales*	*Nitrosomonadaceae*	unclassified	unclassified
Otu00074	10129	*Proteobacteria*	*Alphaproteobacteria*	unclassified	unclassified	unclassified	unclassified
Otu00076	10022	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Oxalobacteraceae*	*Ralstonia*	unclassified
Otu00077	9528	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	ACK-M1	unclassified	unclassified

Otu00082	8022	*Proteobacteria*	*Alphaproteobacteria*	*Rhodospirillales*	*Acetobacteraceae*	*Roseomonas*	*Roseomonas stagni*
Otu00083	8021	*Proteobacteria*	*Betaproteobacteria*	*Nitrosomonadales*	*Nitrosomonadaceae*	unclassified	unclassified
Otu00084	7807	*Firmicutes*	*Bacilli*	*Bacillales*	*Staphylococcaceae*	*Staphylococcus*	unclassified
Otu00085	7773	*Bacteroidetes*	*Cytophagia*	*Cytophagales*	*Cyclobacteriaceae*	*Algoriphagus*	*Algoriphagus aquatilis*
Otu00086	7597	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Oxalobacteraceae*	*Massilia*	unclassified
Otu00089	7368	*Cyanobacteria*	4C0d-2	MLE1-12	unclassified	unclassified	unclassified
Otu00090	7230	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	unclassified	unclassified	unclassified
Otu00092	6903	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Streptococcaceae*	*Streptococcus*	*Streptococcus infantis*
Otu00094	6389	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	unclassified	unclassified
Otu00095	6361	*Proteobacteria*	*Betaproteobacteria*	*Gallionellales*	*Gallionellaceae*	*Gallionella*	unclassified
Otu00096	5934	*Proteobacteria*	*Deltaproteobacteria*	*Desulfobacterales*	*Desulfobulbaceae*	unclassified	unclassified
Otu00097	5867	*Proteobacteria*	unclassified	unclassified	unclassified	unclassified	unclassified
Otu00099	5812	*Proteobacteria*	*Betaproteobacteria*	*Gallionellales*	*Gallionellaceae*	*Gallionella*	unclassified
Otu00100	5684	*Elusimicrobia*	*Elusimicrobia*	*Elusimicrobiales*	unclassified	unclassified	unclassified
Otu00104	5595	*Planctomycetes*	*Planctomycetia*	*Planctomycetales*	*Planctomycetaceae*	*Planctomyces*	unclassified
Otu00106	5197	*Bacteroidetes*	*Flavobacteriia*	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*	unclassified
Otu00109	5004	unclassified	unclassified	unclassified	unclassified	unclassified	unclassified
Otu00110	4964	*Cyanobacteria*	4C0d-2	MLE1-12	unclassified	unclassified	unclassified
Otu00111	4862	*Proteobacteria*	*Alphaproteobacteria*	*Rhodospirillales*	*Acetobacteraceae*	unclassified	unclassified
Otu00115	4625	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	unclassified	unclassified	unclassified
Otu00121	3993	*Chlorobi*	*Ignavibacteria*	*Ignavibacteriales*	*[Melioribacteraceae]*	unclassified	unclassified
Otu00135	3317	*Cyanobacteria*	4C0d-2	MLE1-12	unclassified	unclassified	unclassified
Otu00137	3194	unclassified	unclassified	unclassified	unclassified	unclassified	unclassified

Table A3: Phylum level classification of sequences within each sampling location over the duration of the study shown as relative abundances (%). The phylum *Proteobacteria* is divided into its respective classes

		CHL1	CHL2	CHM1	CHM2	CHM3	CHM4	CHM5	HCHL1	HCHL2	HCHL3
Proteobacteria	*Alphaproteobacteria*	39.24	59.72	38.66	40.78	59.12	52.03	47.99	47.74	56.31	54.00
	Betaproteobacteria	7.33	2.27	2.70	37.74	17.38	28.46	31.63	34.60	25.49	24.57
	Deltaproteobacteria	0.25	0.31	1.30	0.74	1.09	0.50	0.50	0.37	1.20	0.26
	Epsilonproteobacteria	<0.01	<0.01	0.024198	0	0	<0.01	<0.01	<0.01	<0.01	0
	Gammaproteobacteria	10.69	5.30	6.26	5.12	5.33	2.33	3.44	2.82	2.70	2.94
	Unclassified Proteobacteria	1.95	4.19	0.85	0.38	0.24	0.29	0.43	0.23	0.87	0.29
	Planctomycetes	16.01	8.53	32.44	4.82	8.12	6.56	6.22	5.81	4.27	9.38
	Bacteroidetes	1.81	0.75	5.98	2.10	1.83	1.68	2.14	0.92	1.75	1.36
	Cyanobacteria	5.20	6.83	5.23	0.45	1.70	1.50	0.78	0.52	0.76	1.46
	unclassified	6.61	7.94	1.65	1.13	0.91	0.88	0.77	0.51	0.86	0.99
	Nitrospirae	0.17	0.04	0.03	2.86	0.44	0.85	3.30	4.11	2.73	2.03
	Actinobacteria	6.47	1.30	2.11	1.11	1.03	1.82	0.69	0.64	0.85	0.95
	Firmicutes	1.99	2.29	1.77	0.50	0.58	0.55	0.19	0.20	0.48	0.44
	Acidobacteria	1.16	0.14	0.45	0.77	0.73	0.90	0.55	0.41	0.40	0.44
	Crenarchaeota	0.39	0.04	0.25	0.15	0.16	0.31	0.44	0.54	0.07	0.22
	Chlorobi	0.13	<0.01	<0.01	0.52	0.51	0.09	0.14	0.02	0.08	0.12
	Elusimicrobia	<0.01	0.03	0.01	0.20	0.08	0.55	0.07	0.02	0.42	0.15
	Verrucomicrobia	0.21	0.03	0.09	0.24	0.19	0.21	0.17	0.16	0.27	0.12
	Gemmatimonadetes	0.06	<0.01	0.03	0.14	0.10	0.12	0.20	0.07	0.07	0.02
	Chloroflexi	0.04	<0.01	0.03	0.11	0.11	0.05	0.06	0.06	0.13	0.05
	OD1	0.05	<0.01	0.01	0.03	0.04	0.06	0.05	0.04	0.06	0.03
	Fusobacteria	0.08	0.05	0.04	<0.01	0.22	0.06	<0.01	0	0.01	<0.01
	Chlamydiae	<0.01	<0.01	0.02	0.03	0.04	0.06	0.04	0.04	0.04	<0.01
	Spirochaetes	<0.01	0	0.01	0.01	0.04	0.07	0.08	<0.01	0.06	0.07
	OP3	0.06	<0.01	<0.01	0.01	0.07	0.02	0.01	0.01	<0.01	<0.01
	TM6	<0.01	<0.01	0.01	0.01	0.02	0.02	0.03	0.02	0.03	0.02
	NKB19	0	0	0.00	0.01	0.01	<0.01	0.05	0.11	<0.01	0.03
	SBR1093	<0.01	0.15	<0.01	0.00	<0.01	0.02	<0.01	<0.01	0	<0.01
	Armatimonadetes	0.01	0.02	0.01	<0.01	0.01	0.02	0.01	<0.01	<0.01	<0.01
	WPS-2	0	<0.01	0	<0.01	<0.01	0	0.01291	<0.01	0.05	<0.01
	Euryarchaeota	<0.01	<0.01	0.02	<0.01	0.01	0.01	<0.01	<0.01	<0.01	<0.01
	[Thermi]	0.02	<0.01	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	ZB3	<0.01	0	0	<0.01	<0.01	<0.01	0.01	<0.01	<0.01	<0.01
	OP1	0	0	0	<0.01	0	0.03	<0.01	<0.01	<0.01	<0.01
	GN02	0	<0.01	0	<0.01	<0.01	<0.01	0.01	<0.01	<0.01	<0.01
	GN04	<0.01	0	0	<0.01	0.01	<0.01	0	<0.01	<0.01	<0.01
	TM7	0.01	<0.01	0.00	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	PAUC34f	<0.01	0	0	<0.01	<0.01	<0.01	0.01	0	<0.01	<0.01
	[Parvarchaeota]	<0.01	0	0	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	GOUTA4	0	0	0	0	0.01	0	0	0	0	0.01
	WS3	0	0	0	<0.01	<0.01	<0.01	<0.01	<0.01	0.01	<0.01
	NC10	0	0	0	0	0.01	0	0	0	0	0
	Tenericutes	<0.01	0	<0.01	0	<0.01	0	0	0	<0.01	0
	FBP	0	0	0.01	0	0	0	0	0	0	0
	BRC1	0	0	0	0	0	<0.01	0	<0.01	<0.01	<0.01
	Synergistetes	0	0	0	0	0	0	0	<0.01	0	0
	OP11	0	0	0	0	0	0	0	<0.01	0	<0.01
	Lentisphaerae	<0.01	0	0	0	0	0	0	0	<0.01	0

SR1	<0.01	0	0	0	0	0	0	0	0	0
AncK6	<0.01	0	0	0	0	0	0	0	0	0
Aquificae	0	0	0	0	0	0	0	0	0	0
Caldithrix	0	0	0	0	0	0	0	0	0	0
TPD-58	0	0	0	0	0	0	0	0	0	0
Thermotogae	0	0	0	0	0	0	0	0	0	0
Kazan-3B-28	0	0	0	0	0	0	0	0	0	0
Fibrobacteres	0	0	0	0	0	0	0	0	<0.01	0
WWE1	0	0	0	0	0	0	0	0	<0.01	0
WS2	0	0	0	0	0	0	0	<0.01	0	0
[Caldithrix]	0	0	0	0	0	0	0	0	0	0
Caldiserica	0	0	0	0	0	0	0	0	0	0
WS5	0	0	0	0	0	0	0	0	0	0
AD3	0	0	0	0	0	0	0	0	0	0
AC1	0	0	0	0	0	0	0	<0.01	0	0
MVP-21	0	0	0	0	0	0	0	0	0	0
OP8	0	0	0	0	0	0	0	0	0	0
KSB3	0	0	0	0	0	0	0	0	0	0

Table A4: The mean and standard deviation (SD) for temperature, chlorine (free Cl_2 and total Cl_2) and monochloramine residual data for all sample sites averaged for each month of sampling

		Free Cl_2 (mg/l)		Total Cl_2 (mg/l)		Monochloramine (mg/l)		Temperature (°C)	
Season	**Date**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Spring	Oct-14	0.25	0.37	1.51	0.77	1.27	0.83	17.60	1.41
Spring	Nov-14	0.19	0.24	1.52	0.64	1.33	0.74	19.71	1.03
Summer	Dec-14	0.11	0.21	1.61	0.40	1.49	0.52	19.15	3.15
Summer	Jan-15	0.21	0.29	1.16	0.63	0.95	0.76	22.56	1.70
Summer	Feb-15	0.11	0.17	1.14	0.82	1.03	0.86	23.07	0.91
Autumn	Mar-15	0.20	0.30	0.75	0.68	0.55	0.67	21.88	0.92
Autumn	Apr-15	0.18	0.41	0.89	0.79	0.71	0.76	21.45	2.47
Autumn	May-15	0.20	0.40	0.90	0.73	0.70	0.75	18.83	1.61
Winter	Jun-15	0.12	0.25	0.97	0.76	0.85	0.79	17.10	1.81
Winter	Jul-15	0.16	0.26	1.42	0.69	1.26	0.76	13.68	2.27
Winter	Aug-15	0.20	0.42	1.58	0.65	1.39	0.84	12.11	1.44
Spring	Sep-15	0.17	0.22	1.43	0.72	1.26	0.83	15.33	1.54
Spring	Oct-15	0.11	0.26	1.65	0.34	1.53	0.50	18.31	2.40
Spring	Nov-15	0.23	0.21	1.57	0.56	1.34	0.69	19.27	0.67
Summer	Dec-15	0.13	0.20	1.45	0.53	1.32	0.62	21.78	1.59
Summer	Jan-16	0.23	0.43	1.50	0.55	1.27	0.73	23.63	2.00
Summer	Feb-16	0.13	0.27	1.00	0.62	0.87	0.66	24.44	1.37
Autumn	Mar-16	0.14	0.29	1.12	0.70	0.98	0.75	23.64	0.98
Autumn	Apr-16	0.23	0.34	1.26	0.60	1.02	0.78	21.84	0.88
Autumn	May-16	0.33	0.45	1.52	0.39	1.19	0.73	18.29	1.74
Winter	Jun-16	0.22	0.33	1.24	0.69	1.02	0.81	16.76	1.60
Winter	Jul-16	0.42	0.42	1.75	0.52	1.33	0.65	13.62	1.64
Winter	Aug-16	0.31	0.45	1.87	0.22	1.57	0.60	12.63	1.54
Spring	Sep-16	0.26	0.30	1.81	0.34	1.54	0.62	14.25	1.18

Table A5: Summary of PERMANOVA results indicating the influence of temporal and spatial grouping on the overall bacterial community membership and structure within the entire DWDS (CHL1 – HCHL3)

Membership based metrics						
Jaccard distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Season overall	3	6.296	2.09866	6.5431	0.04804	0.001***
Sample location	9	14.017	1.55748	4.8558	0.10695	0.001***
Season overall: Sample location	27	11.965	0.44313	1.3816	0.09129	0.001***
Residuals	308	98.79	0.32075		0.75373	
Total	347	131.068			1	
Unweighted UniFrac distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Season overall	3	3.255	1.08487	4.124	0.03156	0.001***
Sample location	9	11.073	1.23035	4.677	0.10739	0.001***
Season overall: Sample location	27	7.759	0.28735	1.0923	0.07525	0.004**
Residuals	308	81.023	0.26306		0.7858	
Total	347	103.109			1	
Structure based metrics						
Bray-Curtis distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Season overall	3	6.827	2.27566	9.2913	0.06276	0.001***
Sample location	9	16.343	1.81593	7.4142	0.15025	0.001***
Season overall: Sample location	27	10.166	0.3765	1.5372	0.09346	0.001***
Residuals	308	75.437	0.24492		0.69353	
Total	347	108.773			1	
Weighted UniFrac distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Season overall	3	4.973	1.65776	11.9069	0.07271	0.001***
Sample location	9	14.18	1.57551	11.3162	0.20732	0.001***
Season overall: Sample location	27	6.361	0.23559	1.6921	0.093	0.001***
Residuals	308	42.882	0.13923		0.62697	
Total	347	68.396			1	

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table A6: Summary of PERMANOVA results indicating the influence of temporal and spatial grouping on bacterial community membership and structure within the chlorinated section of the DWDS (Section 1) (CHL1 – CHL2)

Membership based metrics						
Jaccard distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	6	2.496	0.416	1.40387	0.28605	0.001***
Sample site	1	0.3082	0.30816	1.03996	0.03532	0.422
Yearly seasons: Sample site	6	1.773	0.29549	0.99721	0.20319	0.499
Residuals	14	4.1485	0.29632		0.47544	
Total	27	8.7256			1	
Unweighted UniFrac distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	6	2.1539	0.35899	1.4456	0.28803	0.001***
Sample site	1	0.3343	0.33433	1.3463	0.04471	0.035*
Yearly seasons: Sample site	6	1.5133	0.25222	1.0157	0.20237	0.424
Residuals	14	3.4766	0.24833		0.4649	
Total	27	7.4782			1	
Structure based metrics						
Bray-Curtis distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	6	3.5591	0.59318	2.1932	0.35445	0.001***
Sample site	1	0.5112	0.5112	1.8901	0.05091	0.022*
Yearly seasons: Sample site	6	2.1844	0.36407	1.3461	0.21755	0.028*
Residuals	14	3.7864	0.27046		0.37709	
Total	27	10.0411			1	
Weighted UniFrac distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	6	2.3552	0.39253	2.7447	0.40548	0.001***
Sample site	1	0.2985	0.29849	2.0871	0.05139	0.048*
Yearly seasons: Sample site	6	1.1526	0.19209	1.3432	0.19843	0.095.
Residuals	14	2.0022	0.14301		0.3447	
Total	27	5.8084			1	

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table A7: Summary of PERMANOVA results indicating the influence of temporal and spatial grouping on bacterial community membership and structure within the chloraminated section of the DWDS (section 2) (CHM1 – CHM5)

Membership based metrics						
Jaccard distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	8	2.697	0.33713	1.0881	0.03196	0.092.
Sample site	12	5.303	0.44195	1.4264	0.06285	0.001***
Yearly seasons: Sample site	90	29.913	0.33236	1.0727	0.35446	0.004**
Residuals	150	46.476	0.30984		0.55073	
Total	260	84.389			1	
Unweighted UniFrac distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	8	2.321	0.29012	1.0904	0.032	0.084.
Sample site	12	4.446	0.37049	1.3925	0.0613	0.001***
Yearly seasons: Sample site	90	25.846	0.28717	1.0793	0.35638	0.004**
Residuals	150	39.91	0.26607		0.55031	
Total	260	72.522			1	
Structure based metrics						
Bray-Curtis distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	8	4.649	0.58114	2.19045	0.06422	0.001***
Sample site	12	2.774	0.23119	0.87141	0.03832	0.805
Yearly seasons: Sample site	90	25.179	0.27977	1.0545	0.34778	0.194
Residuals	150	39.796	0.26531		0.54968	
Total	260	72.398			1	
Weighted UniFrac distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	8	2.974	0.37179	2.41979	0.07175	0.001***
Sample site	12	0.942	0.07848	0.51076	0.02272	0.996
Yearly seasons: Sample site	90	14.492	0.16102	1.048	0.34958	0.296
Residuals	150	23.047	0.15365		0.55595	
Total	260	41.455			1	

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table A8: Summary of PERMANOVA results indicating the influence of temporal and spatial grouping on bacterial community membership and structure within the hypochlorinated section (section 3) (HCHL1 – HCHL3)

Membership based metrics						
Jaccard distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	8	3.3931	0.42414	1.28402	0.17253	0.013*
Sample site	2	1.1463	0.57317	1.73522	0.05829	0.002**
Yearly seasons: Sample site	15	4.2263	0.28176	0.85299	0.2149	0.977
Residuals	33	10.9005	0.33032		0.55427	
Total	58	19.6663			1	
Unweighted UniFrac distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	8	2.4068	0.30086	0.93813	0.13204	0.86
Sample site	2	0.6696	0.3348	1.04396	0.03673	0.341
Yearly seasons: Sample site	15	4.569	0.3046	0.94981	0.25065	0.887
Residuals	33	10.583	0.3207		0.58058	
Total	58	18.2285			1	
Structure based metrics						
Bray-Curtis distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	8	3.8918	0.48647	1.8854	0.22227	0.001***
Sample site	2	2.0915	1.04577	4.0531	0.11945	0.001***
Yearly seasons: Sample site	15	3.0112	0.20075	0.778	0.17198	0.96
Residuals	33	8.5145	0.25802		0.48629	
Total	58	17.509			1	
Weighted UniFrac distance						
	Degrees of freedom	**Sum of squares**	**Mean of squares**	**F.Model**	**R2**	**Pr(>F)**
Yearly seasons	8	1.9022	0.23777	1.7753	0.20321	0.002**
Sample site	2	1.0361	0.51807	3.8681	0.11069	0.001***
Yearly seasons: Sample site	15	2.0028	0.13352	0.9969	0.21395	0.503
Residuals	33	4.4198	0.13393		0.47215	
Total	58	9.3609			1	

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Supplementary figures
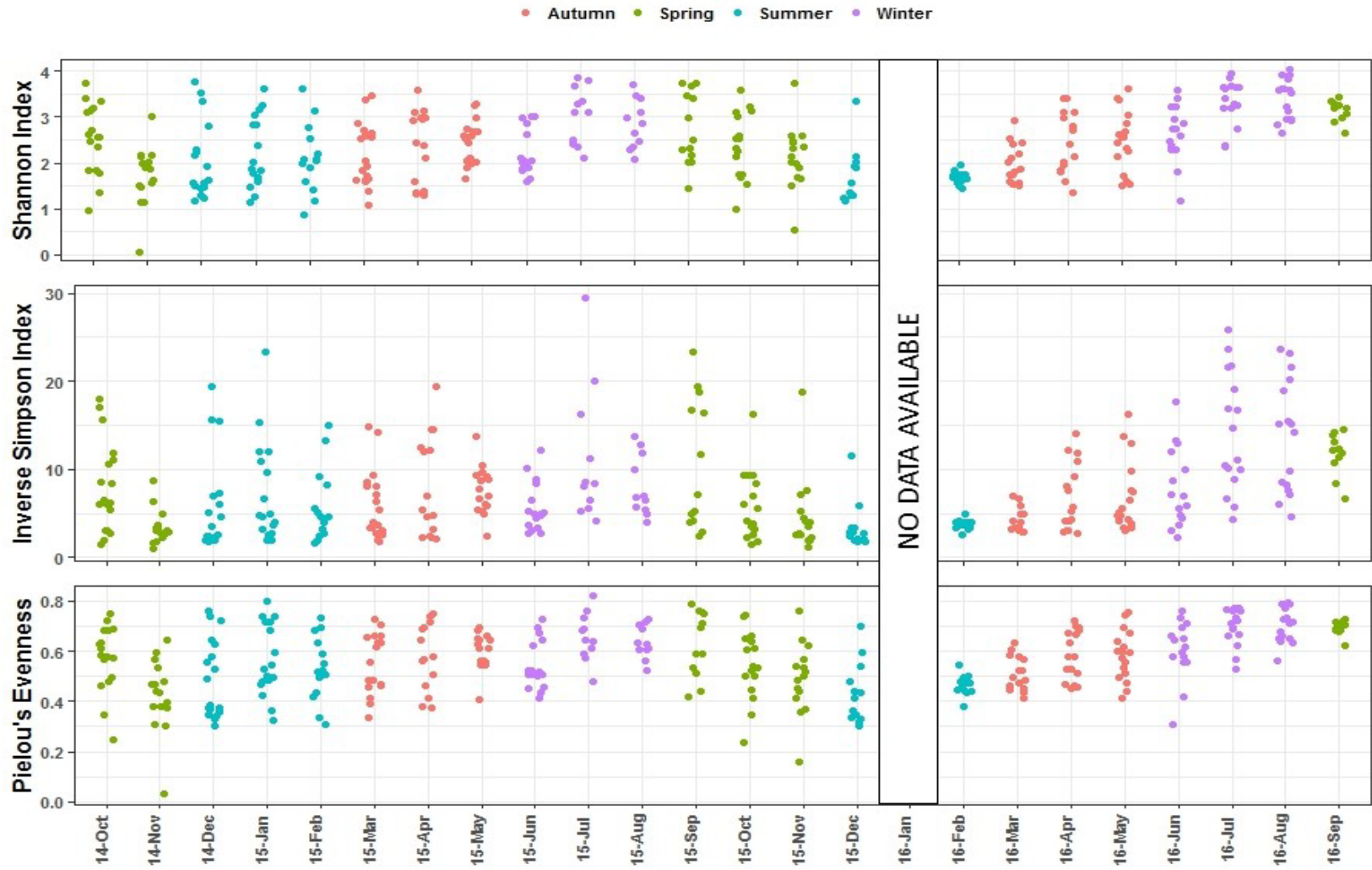
Figure A1: Temporal changes in diversity (Shannon Diversity Index and Inverse Simpson Diversity Index) and evenness (Pielou's evenness) averaged across all sampling locations for each month. Points represent all sample sites collected for each month. Months are coloured based on season indicated in the legend above.
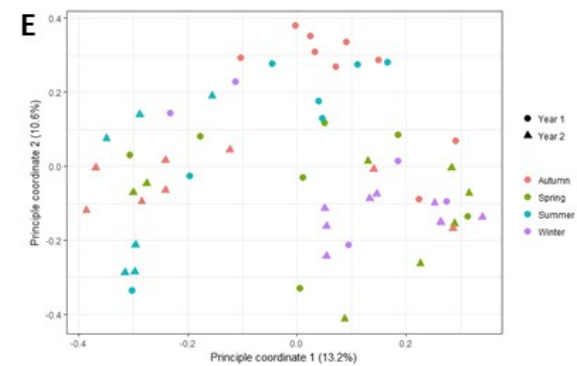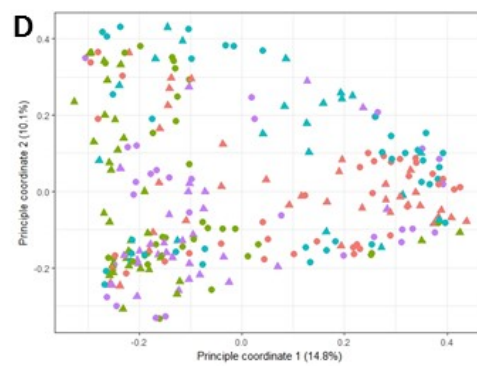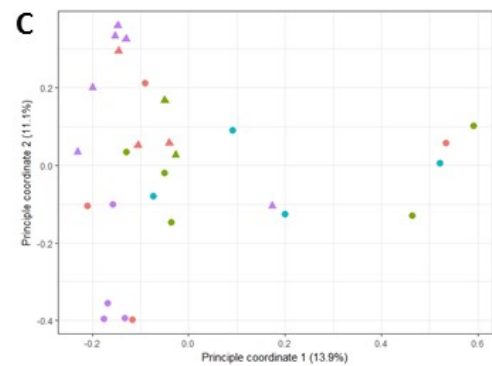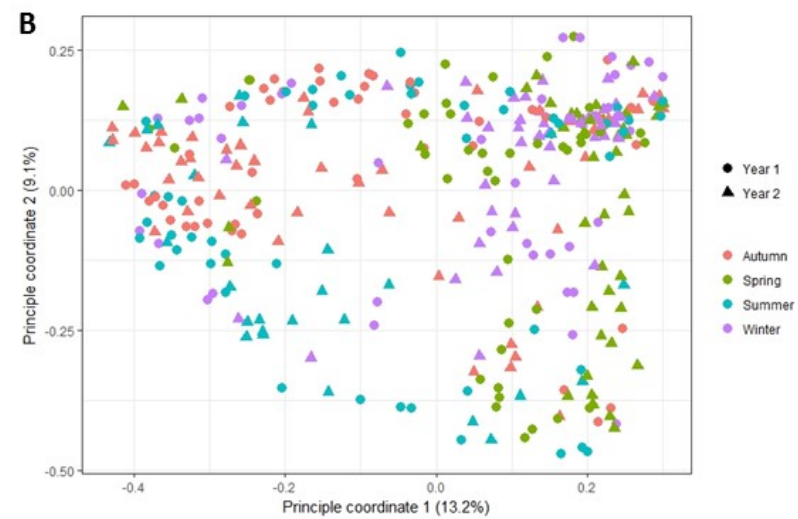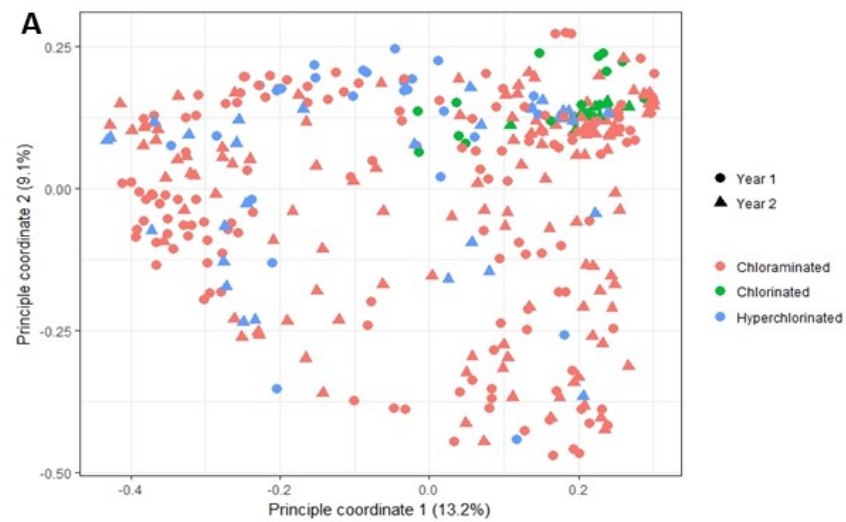
Figure A2: Principal-coordinate biplots showing the spatial and temporal variability of the bacterial community membership within the DWDS using Jaccard dissimilarities distances. Spatial grouping are shown in plot (A) where data points are coloured based on disinfection strategy used and shaped based on year. Temporal grouping are shown in plot (B) where data points are coloured based on season and shaped based on year. Colour and shapes are indicated in the legends on the right of both plots. The three lower plots indicate the temporal groupings within the three sections using different disinfection strategies i.e. chlorinated (C), chloraminated (D) and hyperchlorite (E). These three plots are coloured by season and shaped based on year, shown in the legend on the right of the plot E.
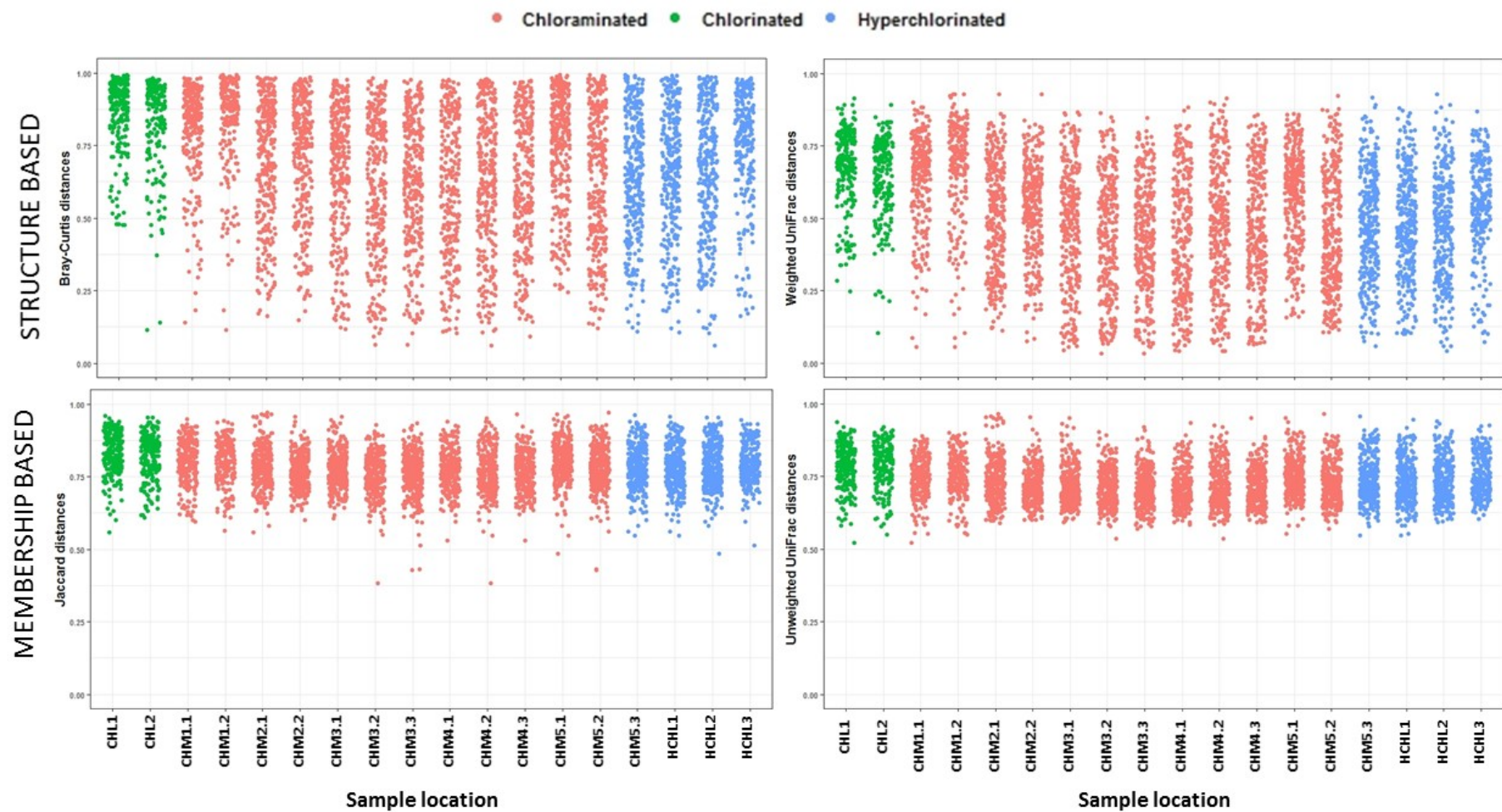
Figure A3: The extent of temporal variation in both the bacterial community structure and membership within each sample site over the duration of the study using pairwise distances of all Beta diversity metrics.
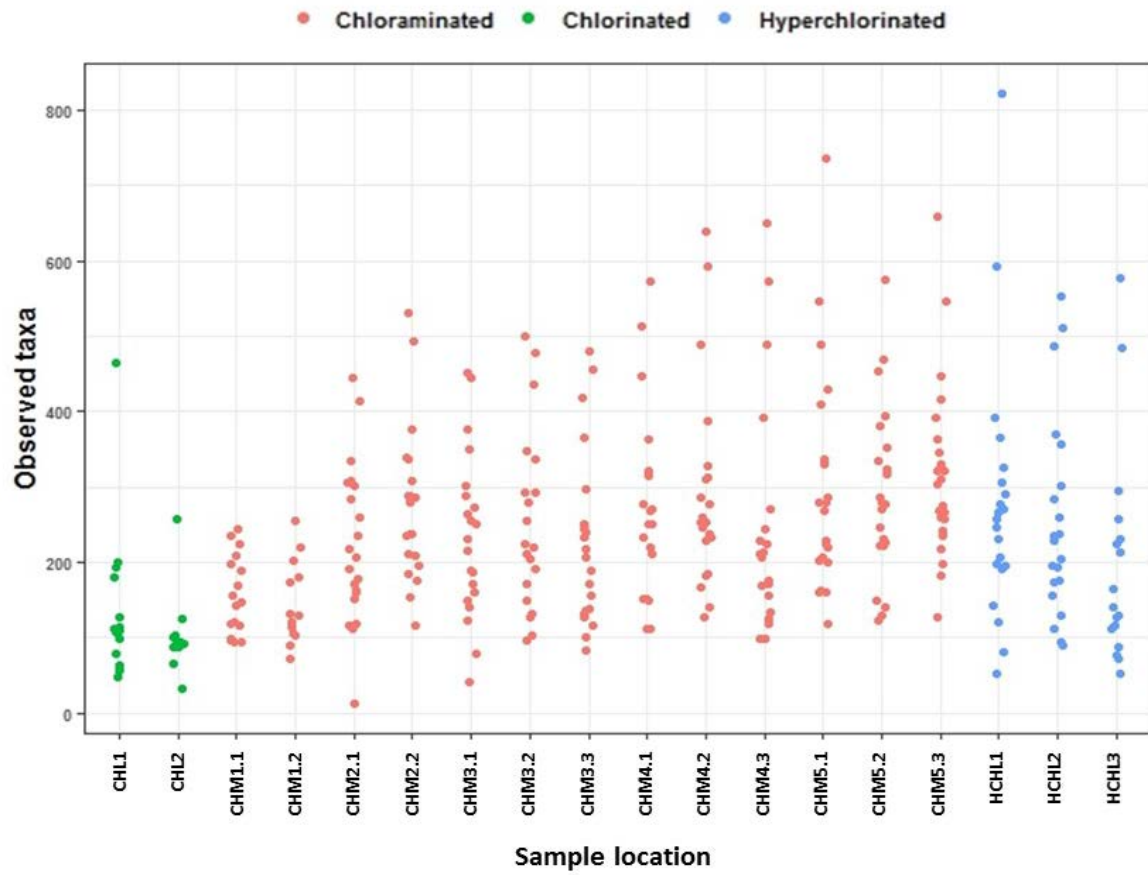
Figure A4: Spatial changes in richness (observed taxa) across the DWDS. Points represent each month over 2-years at each sample site. Sample sites are coloured based on disinfection strategy indicated in the legend on the right.
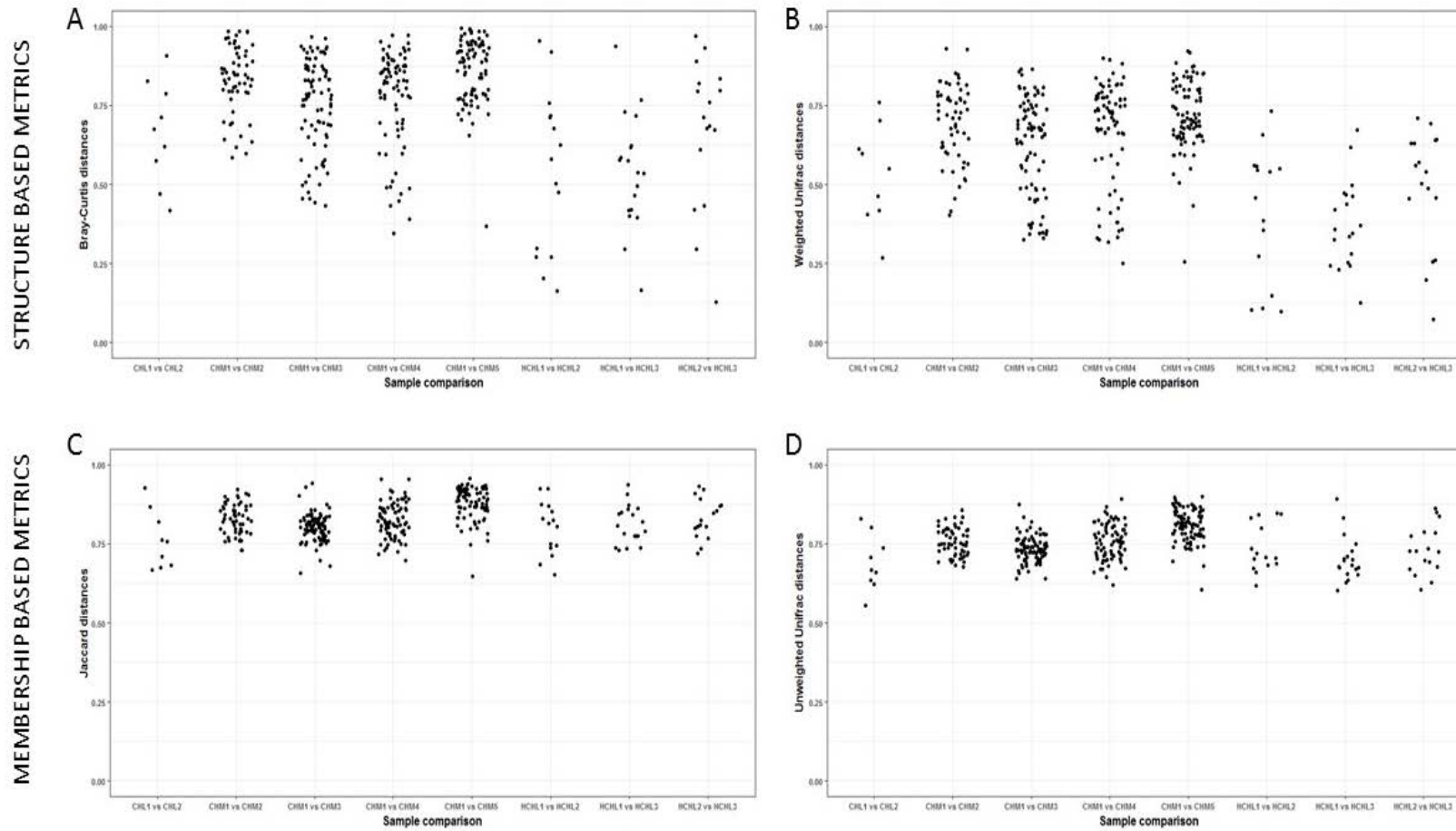
Figure A5: Pair wise beta-diversity distances (structure based metrics: Bray-Curtis [A] and weighted UniFrac [B]; membership based metrics: Jaccard [C] and unweighted UniFrac [D]) between DWDS sample locations within each disinfection section. Comparisons include the first sample after one of the three disinfections compared to all other samples within each disinfection section.

CHAPTER 3

REPRODUCIBLE MICROBIAL COMMUNITY DYNAMICS OF TWO DRINKING WATER SYSTEMS TREATING SIMILAR SOURCE WATERS

Potgieter, S. C., Venter, S. N., Havenga, M., Sigudu, M. and Pinto, A. J. Reproducible microbial community dynamics of two drinking water systems treating similar source waters. (In preparation).

Author Contributions:

SC Potgieter and M Havenga performed all sampling. SC Potgieter performed all analyses and was involved in the conceptualisation and design of the research and the interpretation of the data as well as the drafting of the manuscript. A Pinto was involved in the conceptualisation and design of the research and provided guidance and expertise with regards to sampling processes, data analysis interpretation as well as the drafting of the manuscript. M Sigudu, was part of the Rand Water team involved in this project and was involved in the conceptualisation and design of the project and provided logistical assistance with regards to sampling. SN Venter was involved in conceptualisation and design of the research and the interpretation of the data as well as the drafting of the manuscript.

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.

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 (Figure 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 (Figure 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 B1). 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) (Figure 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 B1, B2A and B2B).

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) (Figure 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) (Figure 1A). 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 (Figure 1B). 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%) (Figure 2A and Table B3).

Between the two varying source waters, only 22.51% of the ASV's identified were shared (i.e., 711 ASV's) (Figure B1). 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 $3.10 \pm 1.51\%$, $3.30 \pm 3.90\%$, $2.95 \pm 0.83\%$ and $2.76 \pm 1.60\%$, 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 $7.63 \pm 1.40\%$, $6.40 \pm 2.55\%$, $4.28 \pm 0.87\%$ and $2.47 \pm 1.61\%$, respectively. Throughout both DWTPs (i.e., including SW, FI, FB and FE samples), *Proteobacteria*, *Actinobacteria* and *Bacteroidetes* were dominant (MRA: $30.51 \pm 10.63\%$ and $26.31 \pm 8.53\%$ and $16.62 \pm 6.83\%$, 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: $9.05 \pm 5.04\%$), *Cyanobacteria* (MRA: $2.89 \pm 2.29\%$), *Verrucomicrobia* (MRA: $2.62 \pm 1.32\%$) and *Planctomycetes* (MRA: $2.07 \pm 2.41\%$) (Figure 2A and Table B3).

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: $6.25 \pm 10.19\%$ and $2.26 \pm 4.44\%$, respectively). This corresponded to increases in *Planctomycetes* and *Cyanobacteria* (MRA: $12.78 \pm 15.32\%$ 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: $17.45 \pm 18.38\%$) and decreased relative abundance of *Proteobacteria* (MRA: $25.50 \pm 7.33\%$), whereas the converse was observed in S_CHLA samples (*Planctomycetes* MRA: $6.24 \pm 6.81\%$ and *Proteobacteria* MRA: $36.51 \pm 15.86\%$). Also, samples from R_CHLB showed higher abundances of *Actinobacteria* and *Bacteroidetes* (MRA: $8.73 \pm 10.30\%$ and $12.30 \pm 18.96\%$, respectively) than S_CHLB samples (*Actinobacteria* MRA: $1.41 \pm 0.06\%$ and *Bacteroidetes* MRA: $0.86 \pm 0.59\%$). Similar to S_CHLA samples, S_CHLB samples also showed an increase in *Proteobacteria* (MRA: $40.03 \pm 18.84\%$) as well as decreased *Planctomycetes* (MRA: $2.77 \pm 4.20\%$) (Figure 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 (Figure 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) (Figure 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%) (Figure 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 (Figure B2, Table B4A and B4B). 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 (Figure 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 (Figure 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 (Figure 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 (Figure 4 and Table B5). 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 (Figure 4 and Table B5). 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 (Figure 4 and Table B5).

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) (Figure 5).

Pairwise beta diversity comparisons between consecutive samples from each system showed similar spatial trends (Figure 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 ≤ 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 (Figure 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 (Figure 7B) and DWDSs (Figure 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 (Figure 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 (Figure 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) (Figure 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) (Figure 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 $\geq$

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 (Figure 8 and Table B6). 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) (Figure 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: $F_{ST} \leq 5.09$, $p < 0.001$) (Figure 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.

6. 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., Garci ́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. *AWWA*. 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
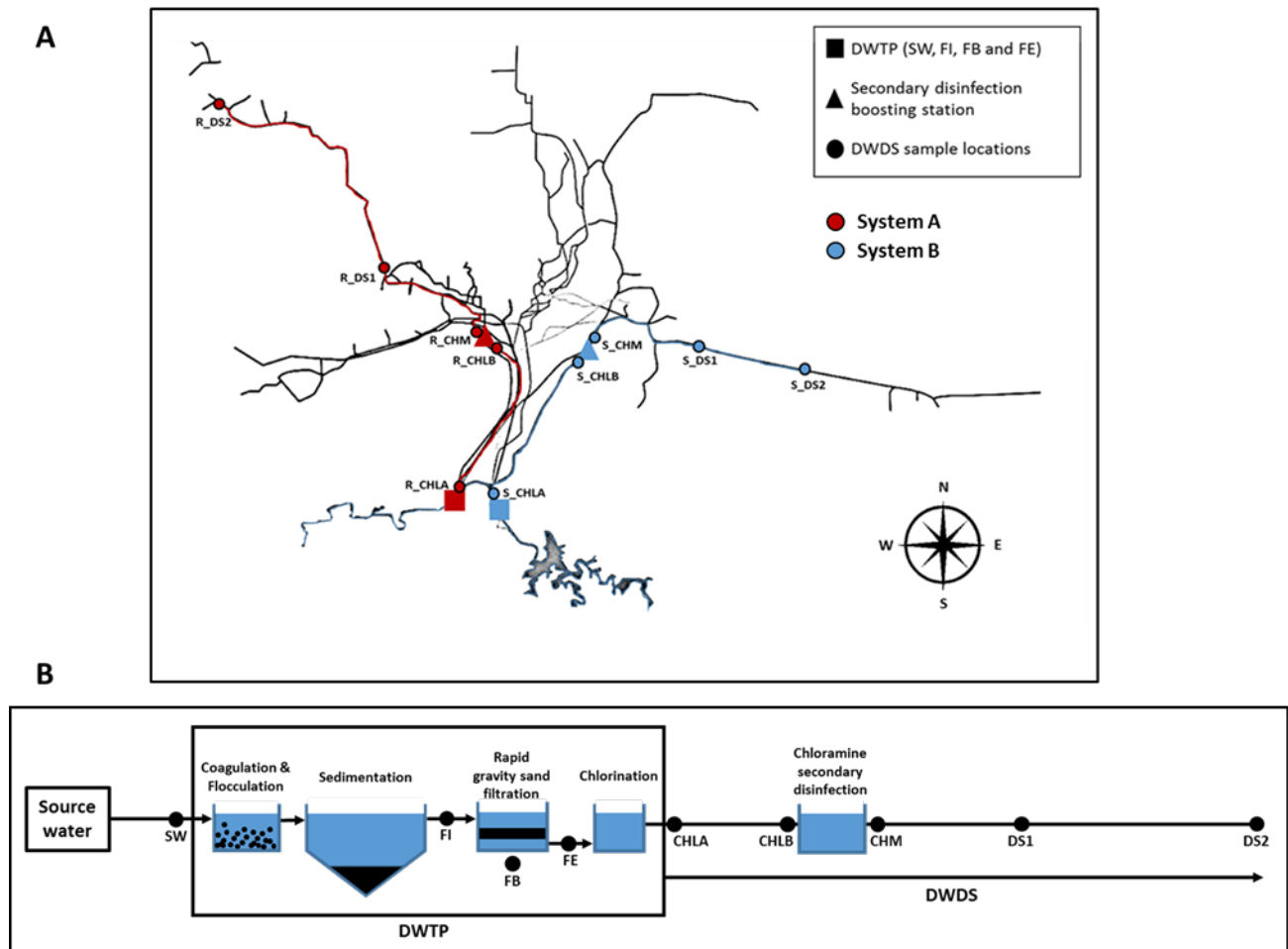
Figure 1: (A) Site map of the location of the drinking water treatment plants (R_DWTP and S_DWTP) and their corresponding distribution systems (R_DWDS and S_DWDS). System R is indicated in red and System S in blue. The two treatment plants are represented as squares, the two-secondary disinfection boosting 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.
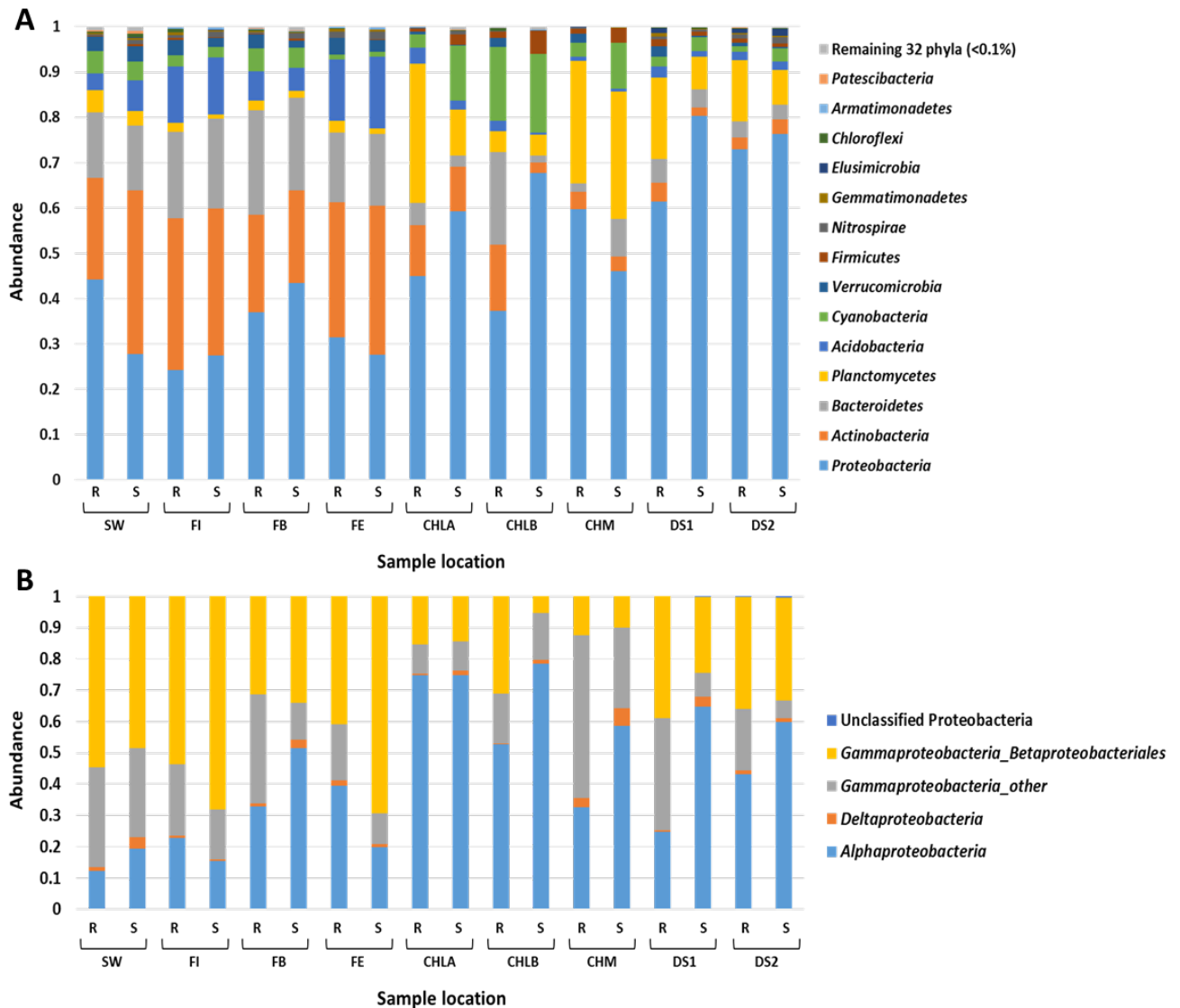
Figure 2: (A) Phylum-level mean relative abundance of bacterial sequences detected over the duration of the study at each sample location within the two DWTPs and corresponding DWDS (R and S DWDS sections). The 14 most abundant and unclassified phyla (> 0.1%) are shown here, with the remaining 32 phyla (< 0.1%) grouped together as a single group. Phyla are shown in the legend on the right of the figure. See Table B1 for mean relative abundances and standard deviations. (B) Mean relative abundance of proteobacterial classes detected over the duration of the study at each sample location for each system.
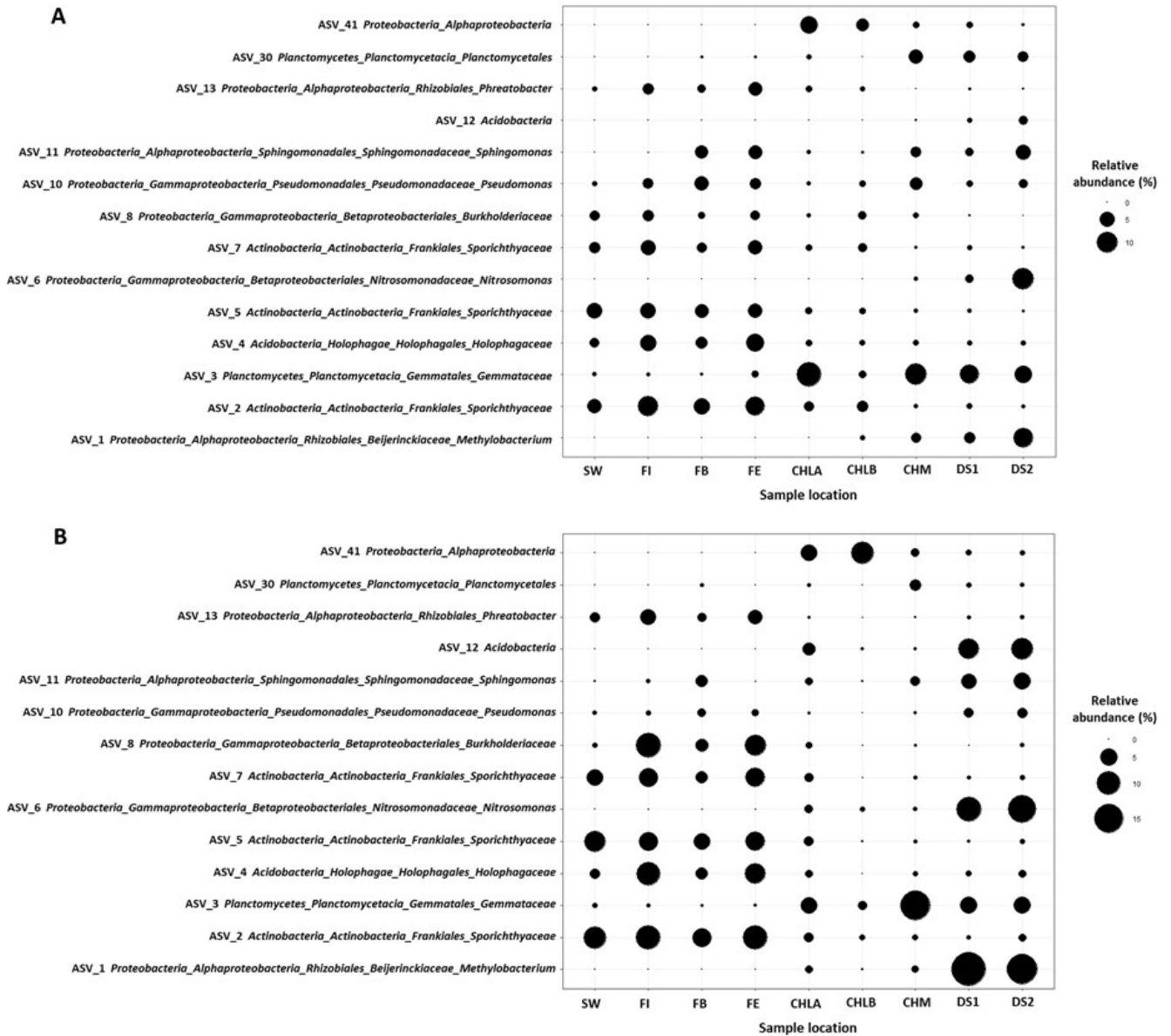
Figure 3: Variation in relative abundance of the 14 most abundant bacterial amplicon sequence variants (ASV's) with a mean relative abundance of ≥ 1% across all samples from (A) System R and (B) System S. The relative abundance for each sample location was averaged over duration of the study for each system. Percentage relative abundance of each ASV is indicated in the legends on the right if the figures. See Table B2A and B2A for mean relative abundances (MRA) of dominant ASVs.
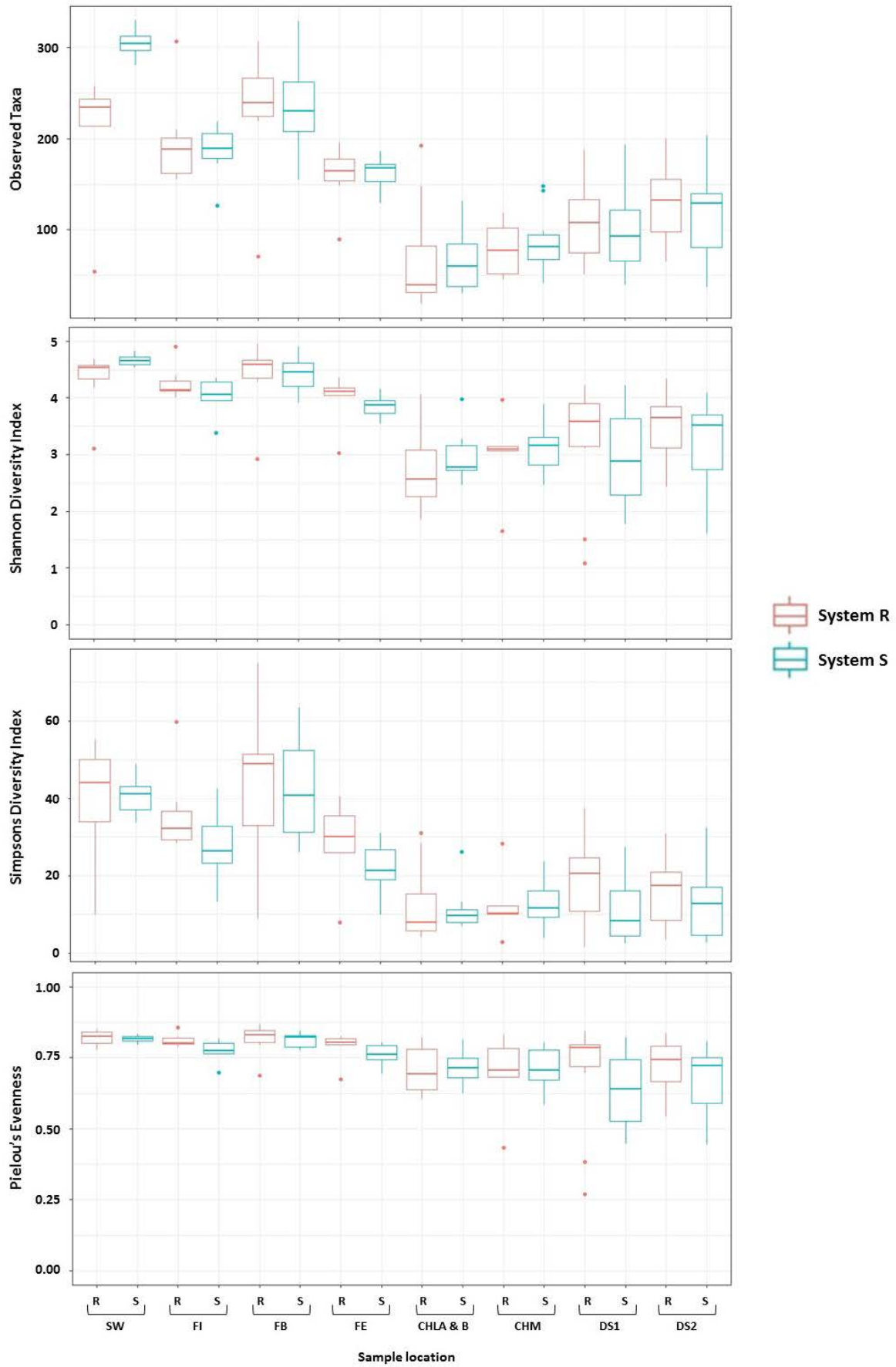
Figure 4: Spatial changes in richness (observed taxa), diversity (Shannon Diversity Index and Inverse Simpson Diversity Index) and evenness (Pielou's evenness) averaged across all sampling locations for each month. Points represent all sample sites collected for each month. Samples coloured based on DWTP and corresponding DWDS (Lines R and S) (subsampled at 1263 iters=1000).
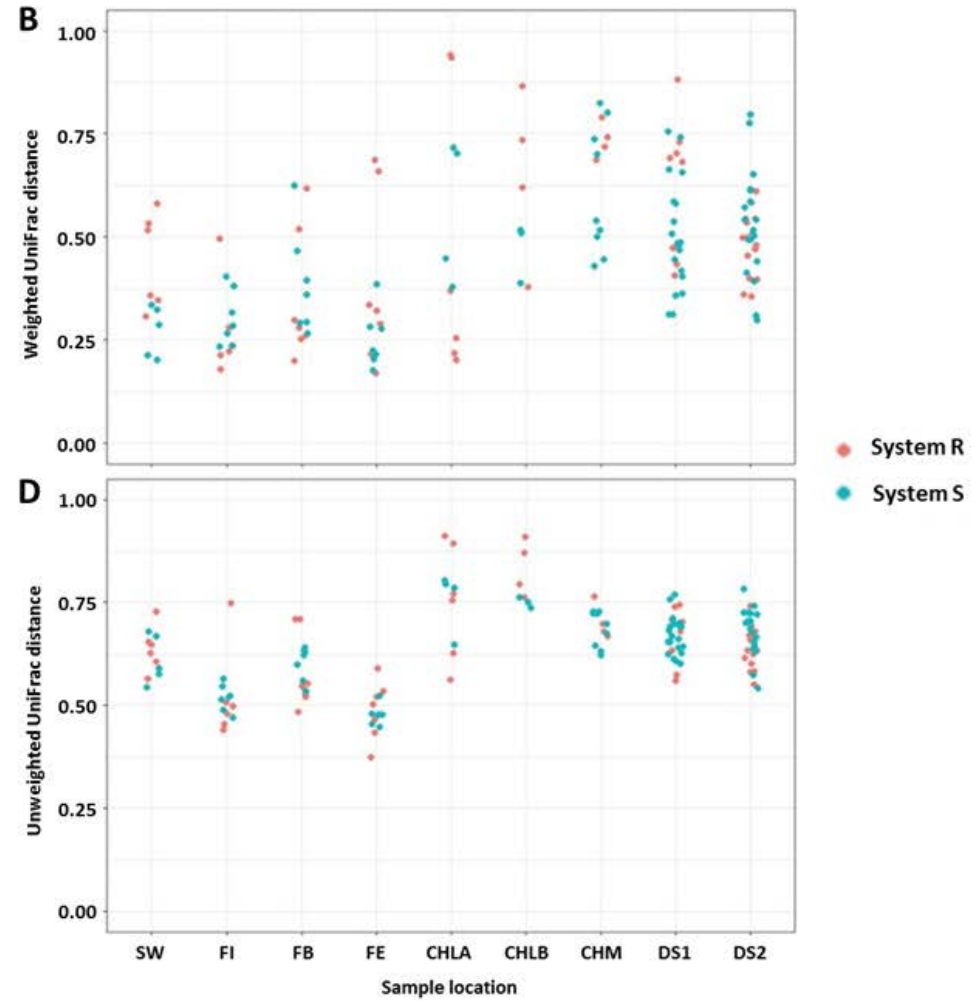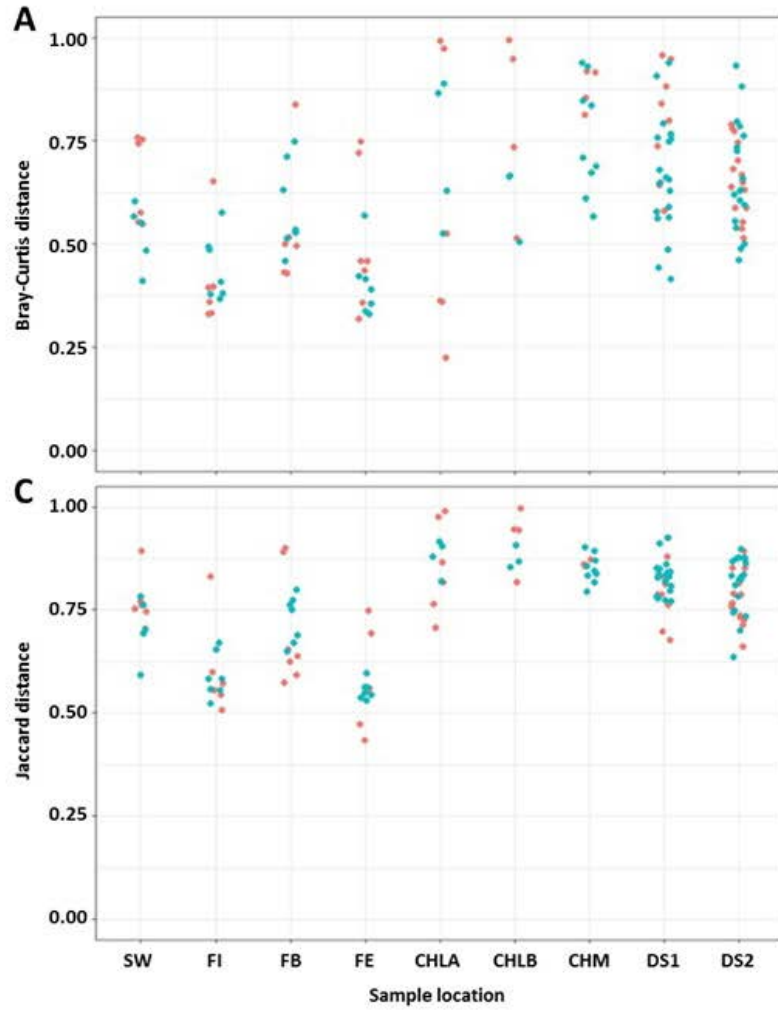
Figure 5: Temporal variation within each sample location. Beta diversity pair-wise comparisons include samples from consecutive months within each location over the eight month study period for both structure based metrics: (A) Bray-Curtis, (B) Weighted UniFrac and membership based metrics: (C) Jaccard, (D) Unweighted UniFrac. Samples form System R are indicated in red and samples from System S are indicated in blue.
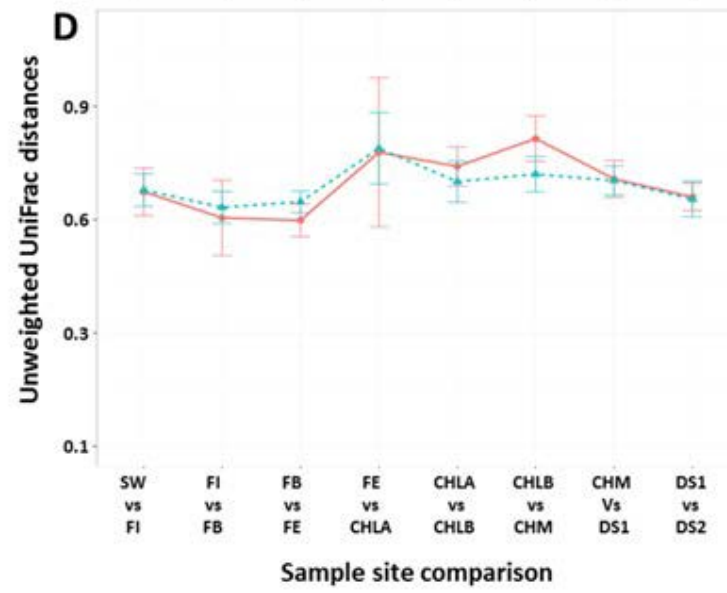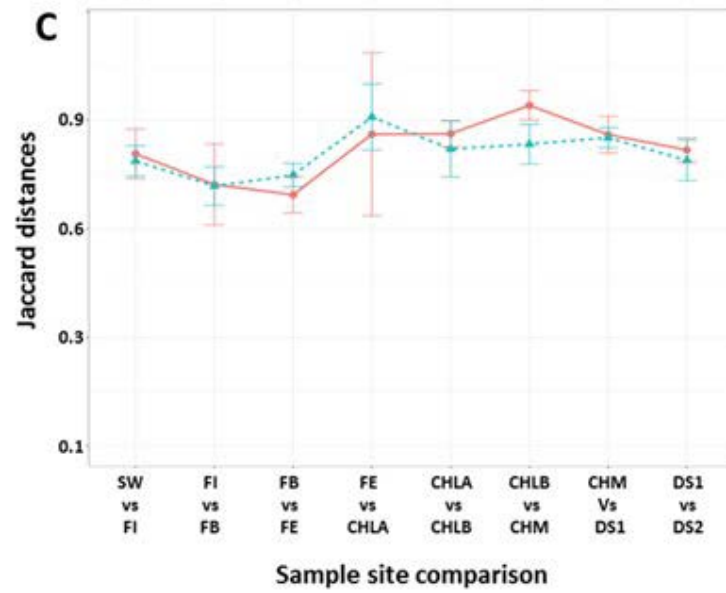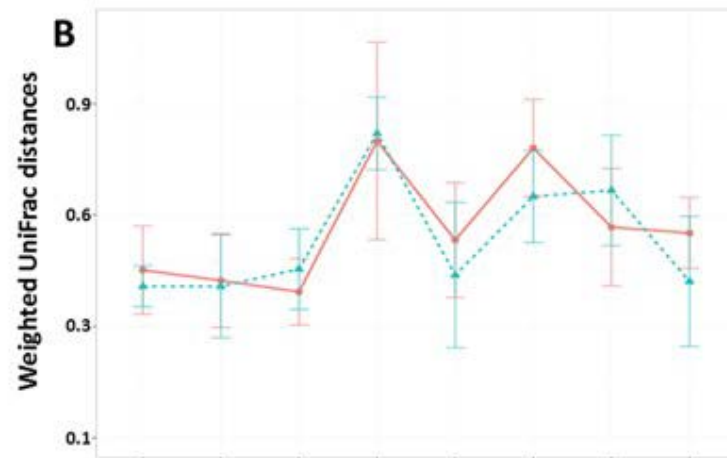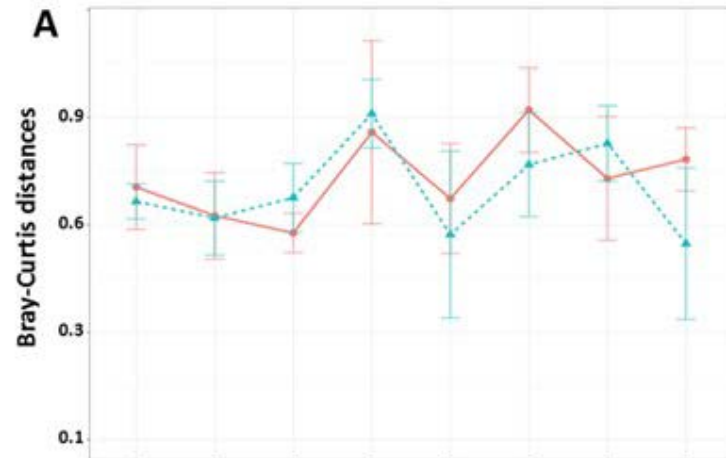
Figure 6: Average pairwise beta diversity comparisons [structure based metrics: (A) Bray-Curtis, (B) Weighted UniFrac and membership based metrics: (C) Jaccard, (D) Unweighted UniFrac] between consecutive locations within each of the two systems for corresponding months. Sample comparisons from System 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.

153
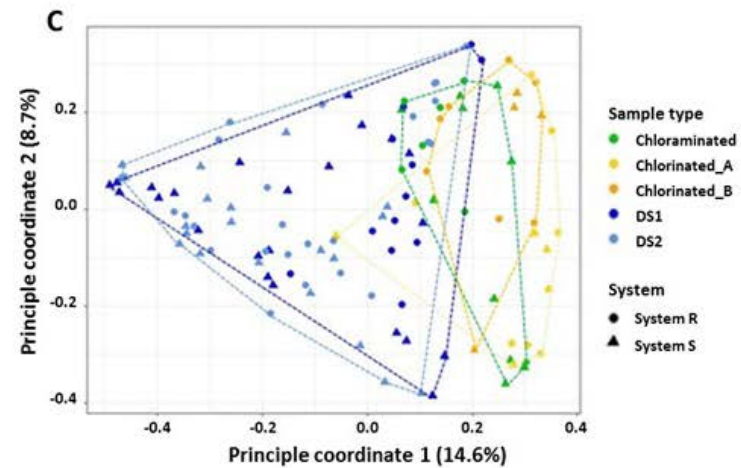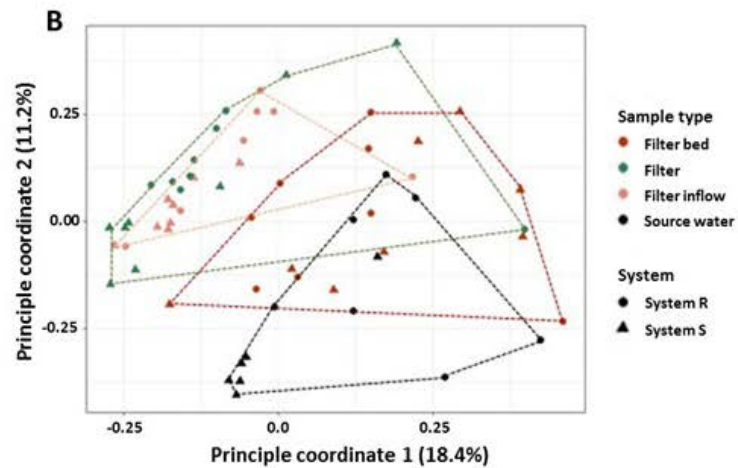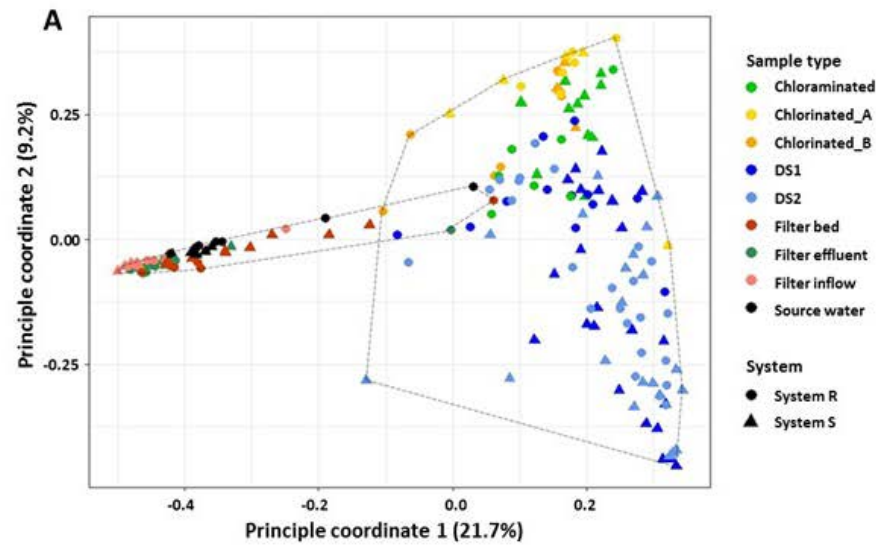
Figure 7: Principal coordinate analysis plot (based on Bray-Curtis dissimilarity) showing the spatial and temporal variability of the bacterial community structure among all samples from both systems (A), within the two DWTPs (B) and within the two corresponding DWDSs (C). Spatial groupings are shown 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.
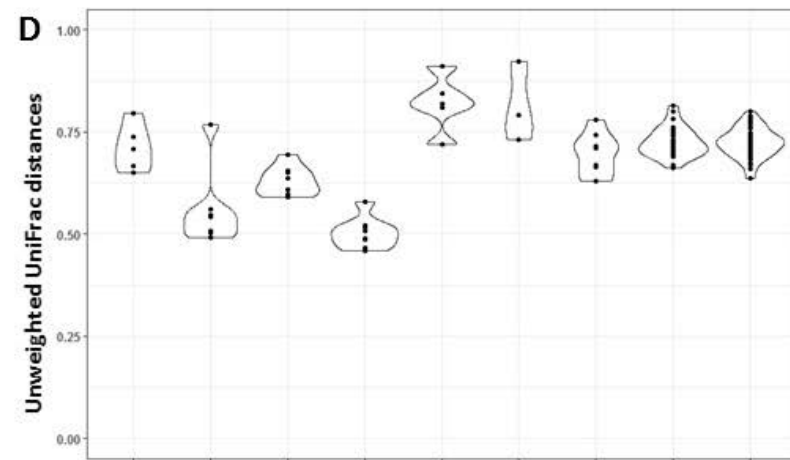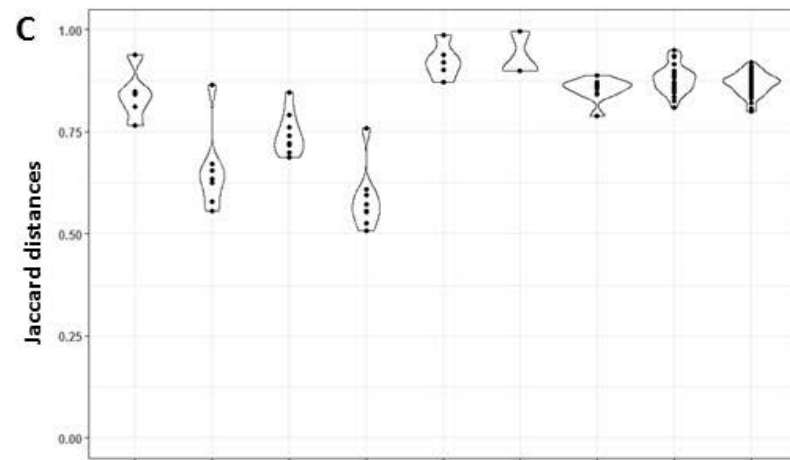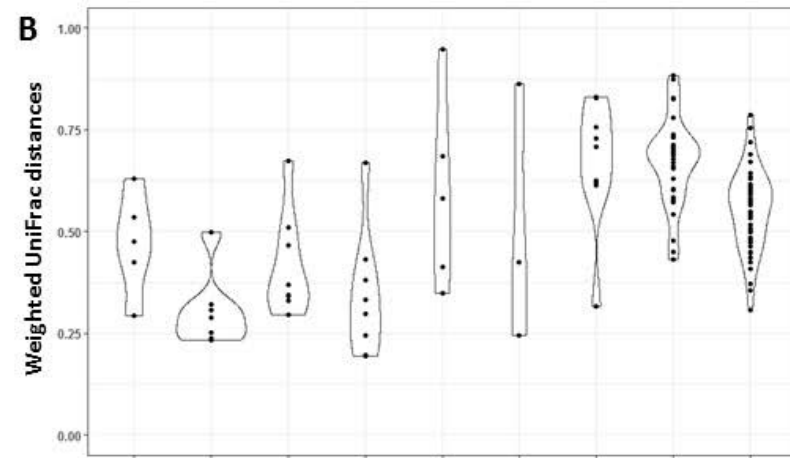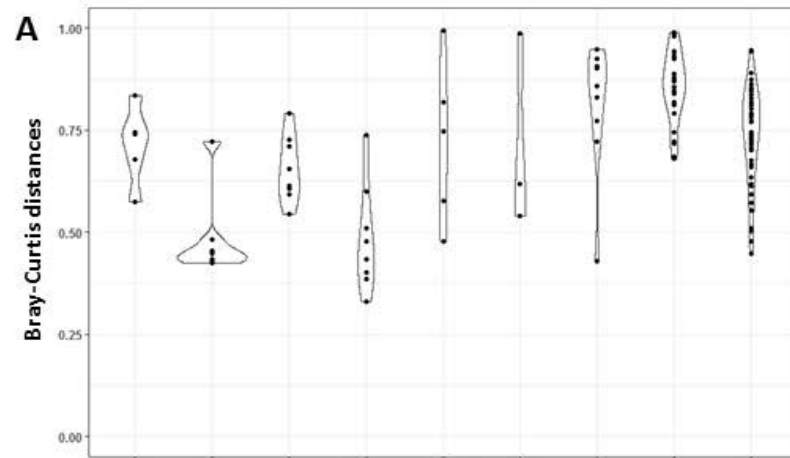
Figure 8: Pairwise beta diversity comparisons between corresponding sample locations from the two systems [structure based metrics: (A) Bray-Curtis, (B) Weighted UniFrac and membership based metrics: (C) Jaccard, (D) Unweighted UniFrac]. Sample abbreviations on the x-axis refer to source water (SW), 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 B3.

Appendix B: Chapter 3 Supplementary information

Supplementary tables

Table B1: Water quality parameters of the source water from both systems

Sample	Date	Source water quality			Treatment process	
		Alkalinity	Turbidity (NTU)	pH	Lime Silica CO2 Ferric	Polymeric Coag lime (low)
System R						
R_SW1	2016/02/22	61	64	8.09	46%	54%
R_SW2	2016/03/22	63	60	7.88	0%	100%
R_SW3	2016/04/19	63	61	8.04	0%	100%
R_SW4	2016/05/17	64	62	8.07	0%	100%
R_SW5	2016/06/20	61	55	8.11	0%	100%
R_SW6	2016/07/18	63	66	8.35	0%	100%
R_SW7	2016/08/16	60	77	8.23	18%	82%
R_SW8	2016/09/19	58	85	8.13	31%	69%
System S						
S_SW1	2016/02/01	58	58	8.2	0%	100%
S_SW2	2016/03/07	67	57	8.3	0%	100%
S_SW3	2016/04/04	59	62	8.3	0%	100%
S_SW4	2016/05/09	58	67	8.5	0%	100%
S_SW5	2016/06/06	59	58	8.7	0%	100%
S_SW6	2016/07/04	57	71	8.5	0%	100%
S_SW7	2016/08/01	57	75	8.8	0%	100%
S_SW8	2016/08/29	51	71	8.7	0%	100%

Table B2A: Water quality parameters for DWDS samples from System R

System R										
Sample	**Month**	**DOC** (mg/l as C)	**Total Cl$_2$** (mg/l)	**Monochloramine** (mg/l)	**NH4** (mg/l as N)	**NO2** (mg/l as N)	**IC_NO3** (mg/l as N)	**pH**	**Temp** (°C)	**Turbidity** (NTU)
R_CHLA1	Feb-16	3.10	2.20	0.00	0.10	0.00	0.10	8.10	24.70	-
R_CHLA2	Mar-16	2.80	2.20	0.00	0.00	0.00	0.10	7.70	21.90	0.33
R_CHLA3	Apr-16	3.40	2.20	0.00	0.00	0.00	0.20	8.00	19.80	0.31
R_CHLA4	May-16	3.20	2.30	0.00	-0.10	0.00	0.10	8.20	16.00	0.42
R_CHLA5	Jun-16	2.60	2.20	0.00	0.00	0.00	0.20	8.00	13.30	0.33
R_CHLA6	Jul-16	0.00	2.20	0.00	0.00	0.00	0.30	8.30	11.50	0.20
R_CHLA7	Aug-16	0.00	1.99	0.00	-0.10	0.00	0.40	8.20	12.20	0.27
R_CHLA8	Sep-16	0.00	2.13	0.00	-0.10	0.00	0.60	8.20	15.00	0.29
R_CHLB1	Feb-16	0.00	1.30	0.00	0.00	0.00	0.00	7.60	24.30	-
R_CHLB2	Mar-16	0.00	1.17	0.00	0.00	0.00	0.00	7.60	22.90	0.35
R_CHLB3	Apr-16	0.00	1.01	0.00	0.00	0.00	0.00	7.90	20.90	0.34
R_CHLB4	May-16	0.00	1.26	0.00	0.00	0.00	0.00	8.10	17.10	0.32
R_CHLB5	Jun-16	0.00	1.27	0.00	0.00	0.00	0.00	8.20	14.70	0.27
R_CHLB6	Jul-16	0.00	1.73	0.00	0.00	0.00	0.20	8.20	12.30	0.36
R_CHLB7	Aug-16	0.00	1.35	0.00	-0.10	0.00	0.20	8.20	12.70	0.34
R_CHLB8	Sep-16	0.00	1.74	0.00	-0.10	0.00	0.20	8.30	15.70	0.14
R_CHM1	Feb-16	0.00	0.06	1.88	0.00	0.00	0.00	7.80	25.20	-
R_CHM2	Mar-16	0.00	0.03	2.00	0.00	0.00	0.00	8.10	22.70	0.28
R_CHM3	Apr-16	0.00	0.07	1.78	0.00	0.00	0.00	8.10	22.00	0.23
R_CHM4	May-16	0.00	0.02	1.77	0.00	0.00	0.00	8.00	17.40	0.22
R_CHM5	Jun-16	0.00	0.04	1.96	0.00	0.00	0.00	8.30	14.40	0.25
R_CHM6	Jul-16	0.00	0.05	1.84	0.10	0.00	0.00	8.40	12.90	0.33
R_CHM7	Aug-16	0.00	0.03	1.63	0.10	0.00	0.00	7.90	13.90	0.30

R_CHM8	Sep-16	0.00	0.10	1.61	0.10	0.00	0.00	8.30	16.40	0.39
R_DS11	Feb-16	0.00	0.07	1.72	0.30	0.05	0.15	8.00	24.45	-
R_DS12	Mar-16	2.35	0.09	1.90	0.40	0.05	0.15	8.00	22.95	0.33
R_DS13	Apr-16	2.75	0.15	1.62	0.40	0.00	0.30	7.75	20.10	0.30
R_DS14	May-16	3.20	0.38	1.61	0.50	0.00	0.15	8.00	17.45	0.35
R_DS15	Jun-16	3.20	0.36	1.56	0.50	0.00	0.20	8.15	14.50	0.30
R_DS16	Jul-16	0.00	0.08	1.67	0.40	0.00	0.20	8.50	12.45	0.19
R_DS17	Aug-16	0.00	0.30	1.46	0.30	0.00	0.35	8.25	12.85	0.28
R_DS18	Sep-16	0.00	0.19	1.66	0.35	0.00	0.25	8.30	16.10	0.32
R_DS21	Feb-16	0.00	0.05	1.14	0.17	0.13	0.10	8.00	27.20	-
R_DS22	Mar-16	2.13	0.39	1.14	0.33	0.07	0.17	7.87	24.07	0.27
R_DS23	Apr-16	1.93	0.39	1.14	0.27	0.07	0.13	8.13	21.30	0.27
R_DS24	May-16	2.17	0.04	1.67	0.27	0.00	0.13	7.97	18.73	0.31
R_DS25	Jun-16	1.60	0.05	1.73	0.47	0.00	0.20	8.40	15.60	0.31
R_DS26	Jul-16	0.00	0.06	1.70	0.43	0.00	0.27	8.30	14.03	0.21
R_DS27	Aug-16	0.00	0.05	1.52	0.20	0.00	0.33	8.13	14.33	0.29
R_DS28	Sep-16	0.00	0.17	1.38	0.33	0.00	0.30	8.30	17.47	0.36

Table B2B: Water quality parameters for DWDS samples from System S

System S										
Sample	**Month**	**DOC (mg/l as C)**	**Total Cl₂ (mg/l)**	**Monochloramine (mg/l)**	**NH4 (mg/l as N)**	**NO2 (mg/l as N)**	**IC_NO3 (mg/l as N)**	**pH**	**Temp (°C)**	**Turbidity (NTU)**
S_CHLA1	Feb-16	0	2.10	0.00	0.00	0.00	0.00	8.25	-	0.22
S_CHLA2	Mar-16	0.00	2.10	0.00	0.00	0.00	0.00	8.25	-	0.18
S_CHLA3	Apr-16	0.00	2.00	0.00	0.00	0.00	0.00	8.60	-	0.11
S_CHLA4	May-16	0.00	1.85	0.00	0.00	0.00	0.00	8.30	-	0.14
S_CHLA5	Jun-16	0.00	2.00	0.00	0.00	0.00	0.00	8.40	-	0.16
S_CHLA6	Jul-16	0.00	2.25	0.00	0.00	0.00	0.00	8.05	-	0.26
S_CHLA7	Aug-16	0.00	2.10	0.00	0.00	0.00	0.00	8.25	-	0.17
S_CHLA8	Sep-16	0.00	1.90	0.00	0.00	0.00	0.00	8.25	-	0.32
S_CHLB1	Feb-16	0.00	1.10	0.02	0.00	0.00	0.00	7.90	24.60	0.30
S_CHLB2	Mar-16	0.00	1.25	0.05	0.00	0.00	0.00	8.10	24.20	0.25
S_CHLB3	Apr-16	0.00	1.25	0.11	0.00	0.00	0.00	8.20	22.00	0.27
S_CHLB4	May-16	0.00	1.22	0.06	0.00	0.00	0.00	8.23	18.40	0.35
S_CHLB5	Jun-16	0.00	0.84	0.07	0.00	0.00	0.00	7.91	16.00	<0.25
S_CHLB6	Jul-16	0.00	1.06	0.13	0.11	0.03	0.15	8.24	11.90	0.15
S_CHLB7	Aug-16	0.00	1.62	0.14	0.05	0.03	0.29	7.85	10.70	0.26
S_CHLB8	Sep-16	0.00	1.04	0.04	0.05	0.03	0.38	7.98	14.70	0.24
S_CHM1	Feb-16	3.50	2.25	2.19	0.35	0.03	0.10	8.31	26.05	0.31
S_CHM2	Mar-16	2.80	1.98	1.93	0.26	0.03	0.10	7.87	21.70	0.25
S_CHM3	Apr-16	3.65	2.17	2.12	0.33	0.03	0.14	8.02	21.30	0.32
S_CHM4	May-16	3.15	2.22	2.14	0.48	0.03	0.14	8.12	18.60	0.33
S_CHM5	Jun-16	3.10	2.16	2.10	0.37	0.03	0.20	8.23	16.05	0.26
S_CHM6	Jul-16	2.20	2.20	2.17	0.19	0.03	0.24	8.24	12.15	0.34
S_CHM7	Aug-16	2.10	2.18	2.12	0.45	0.03	0.28	7.94	10.80	0.30

S_CHM8	Sep-16	0.00	1.17	2.07	0.45	0.03	2.09	8.00	15.65	0.31
S_DS11	Feb-16	3.30	1.17	1.11	0.20	0.11	0.21	7.91	25.53	0.30
S_DS12	Mar-16	2.70	1.79	1.71	0.26	0.03	0.17	7.73	23.77	0.29
S_DS13	Apr-16	3.70	1.66	1.65	0.38	0.03	0.16	8.11	21.43	0.29
S_DS14	May-16	3.10	1.74	1.67	0.53	0.03	0.21	8.04	16.83	0.33
S_DS15	Jun-16	2.60	1.77	1.71	0.46	0.03	0.18	8.30	15.43	0.24
S_DS16	Jul-16	2.30	2.01	0.75	0.50	0.03	0.43	8.07	13.13	0.21
S_DS17	Aug-16	-	2.02	1.71	0.47	0.03	0.30	8.16	12.40	0.33
S_DS18	Sep-16	-	2.07	1.88	0.52	0.03	0.35	7.92	14.90	0.26
S_DS21	Feb-16	3.80	1.00	0.93	0.15	0.17	0.23	8.03	24.90	0.26
S_DS22	Mar-16	2.70	1.39	1.35	0.26	0.03	0.18	7.99	24.07	0.27
S_DS23	Apr-16	3.40	1.50	1.42	0.32	0.03	0.18	8.02	22.13	0.29
S_DS24	May-16	3.40	1.52	1.46	0.46	0.05	0.22	8.15	16.60	0.38
S_DS25	Jun-16	2.30	1.50	1.44	0.43	0.03	0.19	8.27	16.70	0.34
S_DS26	Jul-16	-	1.93	1.35	0.46	0.03	0.36	8.20	13.17	0.23
S_DS27	Aug-16	-	1.97	1.78	0.47	0.03	0.27	8.22	11.77	0.22
S_DS28	Sep-16	-	1.85	1.73	0.51	0.03	0.33	7.95	15.00	0.24

Table B3: Mean relative abundance of dominant bacterial phyla (> 0.1%) across all samples within both DWTPs and their corresponding DWDS lines (R and S)

	SW		FI		FB		FE		CHLA		CHLB		CHM		DS1		DS2	
	R	S	R	S	R	S	R	S	R	S	R	S	R	S	R	S	R	S
Proteobacteria	41.04	23.99	22.63	26.24	33.63	39.19	29.94	26.13	25.50	36.51	22.40	40.02	47.86	36.00	56.03	75.81	68.27	71.90
Actinobacteria	20.69	31.14	31.37	30.94	19.53	18.42	28.27	31.27	6.41	6.01	8.73	1.41	3.07	2.55	3.81	1.79	2.36	3.09
Bacteroidetes	13.49	12.29	17.86	18.96	20.93	18.41	14.66	15.07	2.79	1.53	12.30	0.86	1.40	6.48	4.77	3.84	3.41	3.12
Planctomycetes	4.55	2.77	1.84	0.85	1.94	1.33	2.58	1.12	17.45	6.24	2.74	2.77	21.63	21.98	16.41	6.80	12.59	7.23
Acidobacteria	3.34	5.79	11.57	12.10	5.94	4.62	12.85	14.99	2.04	1.24	1.47	0.24	0.83	0.58	2.17	1.14	1.65	1.68
Cyanobacteria	4.62	3.66	2.31	2.16	4.61	4.01	0.92	1.14	1.59	7.54	9.71	10.30	2.48	7.90	1.95	2.92	1.22	2.81
Verrucomicrobia	3.00	2.95	3.21	1.87	2.76	1.44	3.53	2.46	0.36	0.04	1.27	0.00	1.52	0.02	2.17	0.21	0.65	0.27
Firmicutes	0.13	0.37	0.39	0.21	0.19	0.45	0.05	0.11	0.39	1.46	0.81	2.97	0.92	2.51	1.44	0.85	0.96	0.64
Nitrospirae	0.31	0.71	0.67	1.13	0.25	0.98	1.33	1.64	0.16	0.41	0.07	0.06	0.24	0.03	0.53	0.49	0.88	1.40
Gemmatimonadetes	0.18	0.44	0.54	0.36	0.27	0.17	0.58	0.33	0.03	0.03	0.02	0.01	0.02	0.02	0.75	0.09	0.27	0.22
Elusimicrobia	0.00	0.09	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.06	0.04	0.02	0.91	0.11	0.78	1.51
Chloroflexi	0.36	0.67	0.68	0.14	0.27	0.22	0.11	0.08	0.01	0.05	0.32	0.00	0.00	0.02	0.11	0.19	0.21	0.09
Armatimonadetes	0.05	0.09	0.28	0.31	0.22	0.15	0.24	0.40	0.03	0.00	0.01	0.03	0.00	0.01	0.03	0.01	0.04	0.02
Patescibacteria	0.17	0.48	0.03	0.05	0.09	0.13	0.01	0.02	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.03
Remaining 32 phyla (<0.1)	0.76	0.81	0.16	0.17	0.26	0.64	0.06	0.15	0.01	0.39	0.22	0.40	0.04	0.12	0.14	0.15	0.21	0.28

Table B4A: Mean relative abundance of the most abundance sequence variants (MRA > 1%) across system R samples

	R_SW	R_FI	R_FB	R_FE	R_CHLA	R_CHLB	R_CHM	R_DS1	R_DS2
ASV_1_*Proteobacteria_Alphaproteobacteria_Rhizobiales_Beijerinckiaceae_Methylobacterium*	0.00	0.00	0.00	0.00	0.00	0.63	2.36	2.82	9.14
ASV_2_*Actinobacteria_Actinobacteria_Frankiales_Sporichthyaceae*	4.82	9.41	6.11	8.35	2.29	2.82	0.38	0.65	0.33
ASV_3_*Planctomycetes_Planctomycetacia_Gemmatales_Gemmataceae*	0.45	0.36	0.22	1.04	14.20	1.17	10.63	8.64	6.72
ASV_4_*Acidobacteria_Holophagae_Holophagales_Holophagaceae*	2.01	5.86	3.30	7.11	0.89	0.64	0.79	0.62	0.56
ASV_5_*Actinobacteria_Actinobacteria_Frankiales_Sporichthyaceae*	5.37	5.30	4.38	4.48	1.00	0.90	0.42	0.43	0.14
ASV_6_*Proteobacteria_Gammaproteobacteria_Betaproteobacteriales_Nitrosomonadaceae_Nitrosomonas*	0.01	0.00	0.01	0.00	0.00	0.00	0.49	1.45	10.92
ASV_7_*Actinobacteria_Actinobacteria_Frankiales_Sporichthyaceae*	2.95	5.16	2.23	4.75	0.93	1.77	0.24	0.60	0.23
ASV_8_*Proteobacteria_Gammaproteobacteria_Betaproteobacteriales_Burkholderiaceae_Limnohabitans*	2.18	2.84	1.02	1.92	0.45	1.45	0.77	0.07	0.00
ASV_10_*Proteobacteria_Gammaproteobacteria_Pseudomonadales_Pseudomonadaceae_Pseudomonas*	0.52	2.54	4.45	2.83	0.48	0.87	3.61	0.94	1.69
ASV_11_*Proteobacteria_Alphaproteobacteria_Sphingomonadales_Sphingomonadaceae_Sphingomonads*	0.00	0.05	3.91	4.38	0.48	0.11	2.50	1.59	4.92
ASV_12_*Acidobacteria*	0.62	2.96	1.50	4.12	0.94	0.59	0.02	0.18	0.08
ASV_13_*Proteobacteria_Alphaproteobacteria_Rhizobiales_Phreatobacter*	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.60	1.68
ASV_30_*Planctomycetes_Planctomycetacia_Planctomycetales*	0.00	0.00	0.11	0.11	0.63	0.00	4.59	3.18	2.61
ASV_41_*Proteobacteria_Alphaproteobacteria*	0.01	0.00	0.00	0.01	6.71	3.36	0.81	0.80	0.26

Table B4B: Mean relative abundance of the most abundance sequence variants (MRA > 1%) across system S samples

	S_SW	S_FI	S_FB	S_FE	S_CHLA	S_CHLB	S_CHM	S_DS1	S_DS2
ASV_1_*Proteobacteria_Alphaproteobacteria_Rhizobiales_Beijerinckiaceae_Methylobacterium*	0.01	0.00	0.00	0.01	0.91	0.06	0.74	19.31	15.19
ASV_2_*Actinobacteria_Actinobacteria_Frankiales_Sporichthyaceae*	8.42	9.55	5.74	9.75	1.47	0.56	0.48	0.36	0.86
ASV_3_*Planctomycetes_Planctomycetacia_Gemmatales_Gemmataceae*	0.41	0.25	0.11	0.19	4.20	1.24	14.92	4.68	4.83
ASV_4_*Acidobacteria_Holophagae_Holophagales_Holophagaceae*	1.69	9.44	2.32	7.14	0.87	0.01	0.39	0.55	0.94
ASV_5_*Actinobacteria_Actinobacteria_Frankiales_Sporichthyaceae*	7.63	5.91	4.45	6.34	1.43	0.05	0.27	0.16	0.45
ASV_6_*Proteobacteria_Gammaproteobacteria_Betaproteobacteriales_Nitrosomonadaceae_Nitrosomonas*	0.01	0.03	0.02	0.04	1.10	0.41	0.36	10.32	12.96
ASV_7_*Actinobacteria_Actinobacteria_Frankiales_Sporichthyaceae*	4.28	5.72	2.25	6.25	1.20	0.04	0.22	0.23	0.40
ASV_8_*Proteobacteria_Gammaproteobacteria_Betaproteobacteriales_Burkholderiaceae_Limnohabitans*	0.41	10.47	2.49	7.26	0.64	0.03	0.12	0.03	0.30
ASV_10_*Proteobacteria_Gammaproteobacteria_Pseudomonadales_Pseudomonadaceae_Pseudomonas*	0.32	0.46	1.04	0.85	0.09	0.00	0.14	1.36	1.63
ASV_11_*Proteobacteria_Alphaproteobacteria_Sphingomonadales_Sphingomonadaceae_Sphingomonads*	0.05	0.28	2.32	0.03	0.92	0.06	1.43	3.61	4.70
ASV_12_*Acidobacteria*	1.60	4.13	1.29	3.40	0.13	0.02	0.05	0.22	0.30
ASV_13_*Proteobacteria_Alphaproteobacteria_Rhizobiales_Phreatobacter*	0.00	0.00	0.00	0.00	2.48	0.19	0.14	7.05	7.92
ASV_30_*Planctomycetes_Planctomycetacia_Planctomycetales*	0.00	0.00	0.23	0.00	0.22	0.01	2.00	0.40	0.28
ASV_41_*Proteobacteria_Alphaproteobacteria*	0.00	0.00	0.00	0.00	4.23	8.53	1.02	0.48	0.44

Table B5: Means and standard deviations of the number of sequences and alpha diversity indexes averaged over the duration of the study for each individual study site

	Number of sequences		Number of observed taxa(Sobs)		Inverse Simpson Diversity Index		Shannon Diversity Index		Pielou's evenness		Good's coverage	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
R_SW	34465.71	19422.11	208.52	70.03	39.58	16.05	4.30	0.55	0.82	0.03	0.94	0.03
S_SW	60236.00	17308.25	304.32	17.09	40.69	5.40	4.66	0.10	0.82	0.01	0.88	0.01
R_FI	35137.71	15165.55	196.28	52.47	36.04	11.11	4.27	0.30	0.81	0.02	0.94	0.02
S_FI	49732.13	13430.06	185.92	28.20	27.60	8.75	4.04	0.31	0.77	0.04	0.94	0.01
R_FB	39244.88	21034.37	229.43	70.23	43.79	19.38	4.39	0.62	0.81	0.06	0.92	0.03
S_FB	46060.38	28383.14	232.39	56.01	42.36	13.16	4.41	0.33	0.81	0.03	0.92	0.03
R_FE	35808.13	13983.04	159.10	32.14	28.98	9.96	4.01	0.41	0.79	0.05	0.96	0.01
S_FE	51383.50	24773.13	160.79	20.54	21.84	6.60	3.86	0.20	0.76	0.04	0.95	0.01
R_CHLA	18781.29	16926.91	58.35	42.94	9.94	9.42	2.64	0.67	0.68	0.07	0.99	0.01
S_CHLA	15748.00	11454.43	82.79	36.36	13.08	7.56	3.14	0.55	0.72	0.06	0.99	0.01
R_CHLB	18146.60	28495.57	80.63	76.63	15.65	10.93	3.02	0.97	0.75	0.09	0.98	0.03
S_CHLB	15234.00	25980.55	48.23	25.33	8.86	2.17	2.70	0.17	0.72	0.07	0.99	0.01
R_CHM	10565.40	6432.02	78.96	31.22	12.81	9.36	2.98	0.83	0.69	0.15	0.99	0.01
S_CHM	18030.45	14458.32	86.94	33.26	12.77	6.01	3.14	0.44	0.71	0.07	0.98	0.01
R_DS1	29708.09	25598.37	106.48	41.71	18.13	10.94	3.26	1.04	0.70	0.19	0.98	0.01
S_DS1	23499.52	14113.34	99.37	47.07	10.37	7.19	2.87	0.75	0.63	0.12	0.97	0.02
R_DS2	25236.84	17138.83	129.81	37.40	16.04	7.88	3.50	0.52	0.72	0.08	0.97	0.02
S_DS2	26578.38	16234.82	115.24	45.57	12.99	8.37	3.14	0.80	0.66	0.12	0.97	0.02

Table B6: Pair-wise beta diversity comparisons between corresponding locations from both systems

Sample comparison	Structure based metrics				Membership based metrics			
	Bray-Curtis		Weighted UniFrac		Jaccard		Unweighted UniFrac	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
R_SW vs S_SW	0.71	0.10	0.47	0.13	0.84	0.06	0.71	0.06
R_FI vs S_FI	0.49	0.11	0.31	0.09	0.66	0.10	0.56	0.10
R_FB vs S_FB	0.66	0.08	0.42	0.13	0.75	0.05	0.63	0.03
R_FE vs S_FE	0.48	0.13	0.34	0.16	0.58	0.08	0.50	0.04
R_CHLA vs S_CHLA	0.72	0.20	0.60	0.24	0.92	0.04	0.82	0.07
R_CHLB vs S_CHLB	0.72	0.24	0.51	0.32	0.93	0.06	0.82	0.10
R_CHM vs S_CHM	0.74	0.21	0.67	0.16	0.82	0.06	0.69	0.05
R_DS1 vs S_DS1	0.72	0.23	0.67	0.11	0.83	0.08	0.72	0.04
R_DS2 vs S_DS2	0.67	0.17	0.56	0.10	0.83	0.06	0.72	0.04

Supplementary figures
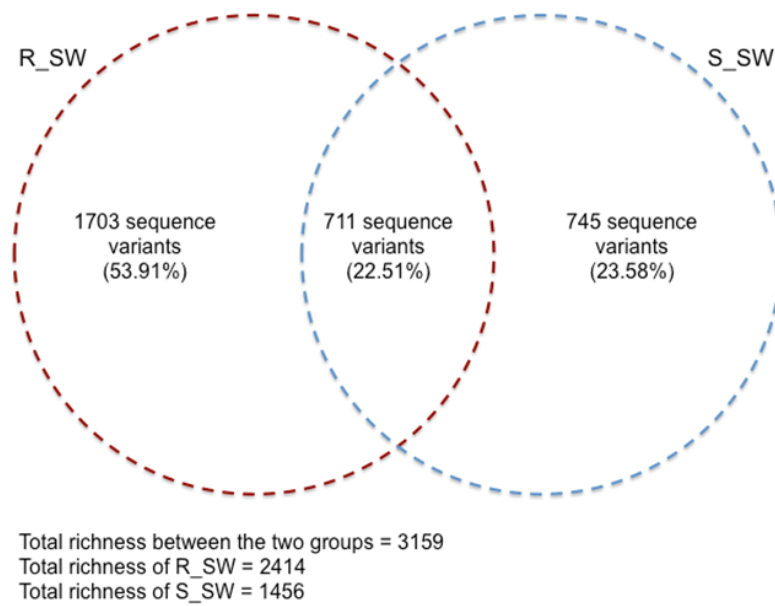
Figure B1: Venn diagram showing the shared amplicon sequence variants (ASVs) between the two source waters (R_SW and S_SW).
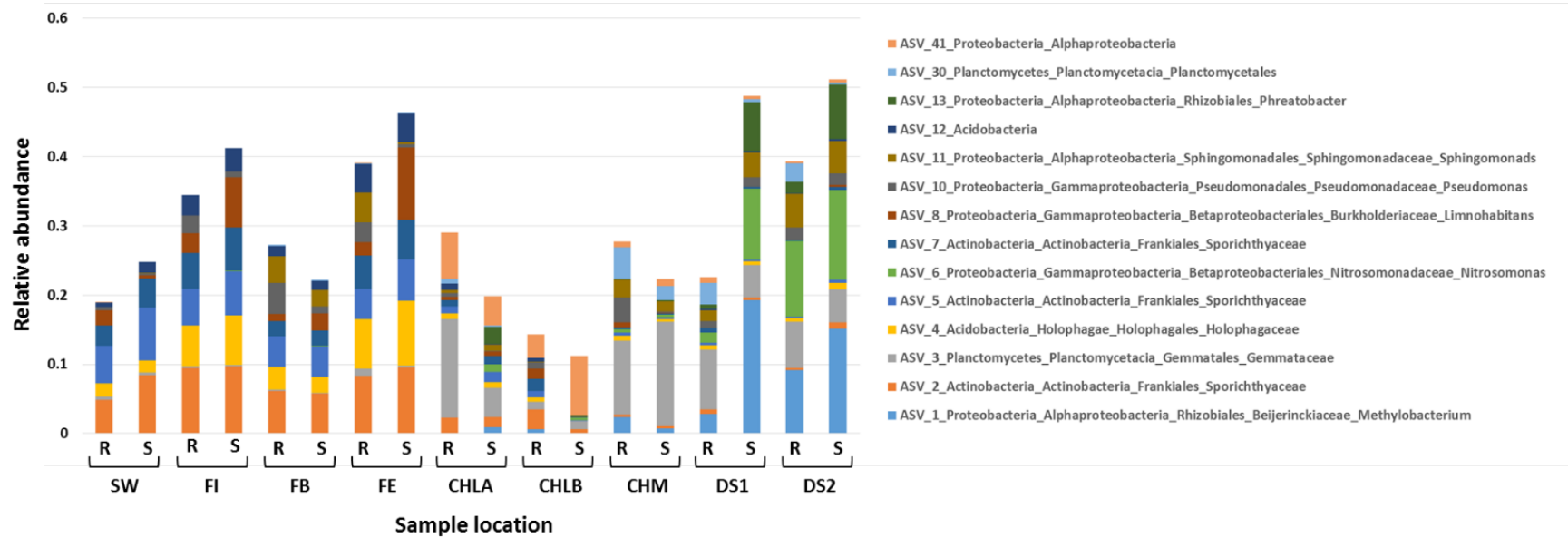
Figure B2: Mean relative abundance of the 14 most abundant bacterial amplicon sequence variants (ASV's) (with a MRA > 1% across all samples) detected over the duration of the study at each sample location within the two DWTPs and corresponding DWDS (R and S DWDS sections). See Tables S2A and S2B for mean relative abundances.

CHAPTER 4

MICROBIAL NITROGEN METABOLISM IN CHLORAMINATED DRINKING WATER RESERVOIRS

Potgieter, S. C., Dai, Z., Venter, S. N., Sigudu, M. and Pinto, A. J. Microbial Nitrogen Metabolism in Chloraminated Drinking Water Reservoirs. (In preparation).

Author Contributions:

SC Potgieter performed all sampling and was involved in the conceptualisation and design of the research and the interpretation of the data as well as the drafting of the manuscript. Z Dai and A Pinto were involved in all sequence processing and data analyses. A Pinto was also involved in the conceptualisation and design of the research and provided guidance and expertise with regards to sampling processes, data analysis interpretation as well as the drafting of the manuscript. M Sigudu, was part of the Rand Water team involved in this project and was involved in the conceptualisation and design of the project and provided logistical assistance with regards to sampling. SN Venter was involved in conceptualisation and design of the research and the interpretation of the data as well as the drafting of the manuscript.

Abstract

Nitrification is often a concern in chloraminated drinking water distribution systems. The addition of ammonia promotes the growth of nitrifying organisms, causing the depletion of chloramine residuals and resulting in operational problems for many drinking water utilities. Therefore, a comprehensive understanding of the microbially mediated processes behind nitrogen metabolism together with chemical water quality data, may allow water utilities to better address the undesirable effects caused by nitrification. In this study, a metagenomic approach was applied to characterise the microbial nitrogen metabolism within chloraminated drinking water reservoirs. Samples from two geographically separated but connected chloraminated reservoirs within the same drinking water distribution system (DWDS) were collected within a 2-year sampling campaign. Spatial changes in the nitrogen compounds (ammonium (NH_4^+), nitrites (NO_2^-) and nitrates (NO_3^-)) across the DWDS were observed, where nitrate concentrations increased as the distance from the site of chloramination increased. The observed dominance of *Nitrosomonas* and *Nitrospira*-like bacteria, together with the changes in the concentration of nitrogen species, suggests that these bacteria play a significant role in contributing to varying stages of nitrification in both reservoirs. Functionally annotated protein sequences were mined for the genes associated with nitrogen metabolism and the community gene catalogue contained mostly genes involved in nitrification, nitrate and nitrite reduction and nitric oxide reduction. Furthermore, based on the construction of Metagenome Assembled Genomes (MAGs), a highly diverse assemblage of bacteria (i.e., predominately *Alpha-* and *Betaproteobacteria* in this study) was observed among the draft genomes. Specifically, 5 MAGs showed high coverage across all samples including two *Nitrosomonas, Nitrospira, Sphingomonas* and a *Rhizobiales*-like MAGs. The role of these MAGs in nitrogen metabolism revealed that the fate nitrate may be linked to changes in ammonia concentrations, that is, when ammonia concentrations are low, nitrate may be assimilated back to ammonia for growth. Alternatively, nitrate may be reduced to nitric oxide and potentially used in the regulation of biofilm formation. Therefore, this study provides insight into the genetic network behind microbially mediated nitrogen metabolism and together with the water chemistry data improves our understanding nitrification in chloraminated DWDSs.

1. Introduction

Disinfection of the drinking water in some DWDSs is often considered key in the management of microbial growth and the maintenance of water quality in most parts of the world, with the exception of a few countries in Europe where microbial growth in drinking water distribution systems (DWDSs) can be managed through nutrient limitation (van der kooij *et al*., 2002; Hammes *et al*., 2008). Chlorine and chloramine have long been successfully used to control microbial growth within DWDSs and although chlorination (primary disinfectant) is successful at initially reducing bacterial growth, distribution system management often now includes chloramination as a secondary disinfectant. Chloramines are typically used to provide disinfectant residuals when free chlorine residuals are difficult to maintain. Chloramines show greater stability as compared to chlorine in the DWDS over long distances, increased efficiency in reducing biofilm growth, and they also produce lower concentrations of regulated disinfection by-products (Norton and LeChevallier, 1997; Vikesland *et al*., 2001; Regan *et al*., 2003; Zhang *et al*., 2009).

However in chloraminated systems, the introduction of ammonia provides an alternative source of nitrogen and growth substrate for ammonia-oxidising microorganisms (AOM), either due to the presence of excess free ammonia or through ammonia released due to chloramine decay (Regan *et al.,* 2003; Zhang *et al*., 2009). This promotes the growth of nitrifying bacteria and archaea, leading to increased nitrification (Belser, 1976; Nicol and Schleper, 2006). Nitrification is an essential process in the biogeochemical nitrogen cycle and links the aerobic and anaerobic pathways of the nitrogen cycle by delivering nitrite and nitrate as electron acceptors for dissimilatory nitrate reduction, denitrification, respiratory ammonification, and anaerobic ammonia oxidation (Kraft *et al.,* 2014; Koch *et al*., 2015). Traditionally microbial nitrification was considered as a two-step process: firstly, ammonium (NH_4^+) is oxidised to nitrite (NO_2^-) by chemolithoautotrophic ammonia-oxidizing bacteria and archaea (AOB and AOA, respectively) (van der Wielen *et al*., 2009) and secondly nitrite is oxidised to nitrate (NO_3^-) by chemolithoautotrophic nitrite-oxidising bacteria (NOB) (Wolfe *et al.,* 1990; Cunliffe, 1991; Francis *et al.*, 2005). NOB are often the principle biological source of nitrate, which is not only an important source of nitrogen for other microorganisms but can also serve as an electron acceptor in the absence of oxygen. In contrast to AOB/AOA and NOB, complete ammonia oxidising bacteria (i.e., comammox) can completely oxidise ammonia to nitrate (Daims, *et al*., 2015; Pinto *et al*., 2015; van Kessel *et*

al., 2015). Unlike AOB, AOA, and NOB, which are phylogenetically diverse, all known comammox bacteria, belong to the genus *Nitrospira* (Phylum: Nitrospirota) (Daims, *et al*., 2015; Pinto *et al*., 2015; van Kessel *et al*., 2015; Fowler *et al*., 2018; Palomo *et al.,* 2018). Furthermore, reciprocal feeding has been described where NOB from the genus *Nitrospira* initiate nitrification by supplying ammonia oxidisers lacking urease and/or cyanase with ammonia from urea or cyanate (Lücker *et* al., 2010; Koch *et al.,* 2015; Palatinszky *et al*., 2015).

Bacterial nitrification in the DWDS causes depletion of chloramine residuals and disinfection decay. The resulting formation of nitrite in the system is problematic as it can rapidly decrease free chlorine and is also further oxidised leading to an accelerated decrease in residual chloramine (Wolfe *et al.,* 1990; Cunliffe, 1991). In addition, due to its toxicity, the regulated concentration of nitrite is typically very low. While ammonia, nitrites and nitrates can serve as an energy source for AOB and NOB (Kirmeyer *et al.,* 1995; Pintar and Slawson, 2003), the loss of chloramine residuals can also lead to heterotrophic bacterial growth and biofilm accumulation, which potentially causes operational problems for many drinking water utilities (Kirmeyer *et al.,* 1995; Norton and LeChevallier, 1997; Pintar and Slawson, 2003).

In a previous study by Potgieter *et al*. (2018), *Nitrosomonas* spp. were observed to be dominant in the chloraminated sections of the DWDS, suggesting that ammonia oxidation and potentially nitrification may be important processes in this DWDS. Therefore, the principle goal of this study was to understand the metabolic potential of microbial communities that might impact the fate of nitrogen in a chloraminated DWDS. Here, the organisms and genes involved in the nitrogen cycle, using a genome resolved metagenomics approach, was investigated. The use of shotgun metagenomic sequencing allowed us to: (i) overcome primer bias, (ii) assemble large operons through *de novo* assembly and (iii) to pinpoint functions to genomes through binning, which can then be phylogenetically identified. Therefore, using this approach, the study aims to explore nitrogen metabolism in chloraminated drinking water reservoirs by (i) investigating the taxonomic profile of the microbial community and the specific genes involved in nitrogen metabolism, (ii) identifying the processes that can drive nitrogen transformation in chloraminated drinking water, and (iii) identifying the role of dominant reconstructed draft genomes in nitrogen metabolism.

2. Materials and methodology

2.1 Site description and sample collection

Sampling was conducted at two geographically separated but connected chloraminated reservoirs within a large South African DWDS previously described by Potgieter *et al.* (2018). Briefly, the process for treating surface water includes coagulation with polymeric coagulants, flocculation, sedimentation, pH adjustment with CO_2 gas followed by filtration (rapid gravity sand filters) and finally initial disinfection with chlorine. Filter effluent is dosed with chlorine to achieve total residual chlorine concentrations varying between 1 and 1.5 mg/L at the outlet of the drinking water treatment plant (DWTP). Chlorinated drinking water is then dosed with chloramine (0.8 to 1.5 mg/L) at a secondary disinfection boosting station approximately 23 km from the DWTP. Here, monochloramine residuals vary seasonally between 0.8 and 1.5mg/L. Within the chloraminated section of the DWDS, the first of the two reservoirs (RES1) sampled is located approximately 32 km from the secondary disinfectant boosting station. The second reservoir (RES2) is located approximately a further 88 km downstream from the first reservoir (Figure 1). Samples were collected within 2 years (October 2014 to September 2016). Further details on a range of chemical parameters, including temperature, disinfectant residual concentrations (i.e., free chlorine, total chlorine, and monochloramine) and nitrogen species concentrations (i.e., ammonium, nitrite and nitrate) were obtained from the utility (Table C1A and C1B).

2.2 Sample processing

Bulk water samples were collected in 8L sterile Nalgene polycarbonate bottles and transported to the laboratory on ice where they were kept at 4°C for 24 to 48 hours until further processing. Samples were filtered to harvest microbial cells by pumping the collected bulk water through STERIVEX™ GP 0.22 µm filter units (Millipore) using a Gilson® minipuls 3 peristaltic pump. The filters were kept in the dark and stored at -20°C until processing and DNA extraction. A traditional phenol/chloroform extraction method optimised by Pinto *et al.* (2012) modified from Urakawa *et al.* (2010) was used for the isolation of DNA from cells immobilised on filter membranes. Following extraction, 8 samples from RES1 and 10 samples from RES2 were selected for shotgun metagenomic sequencing (Table C2).

2.3 Metagenomic sequence processing, *de novo* assembly, functional annotation, and reference mapping

Paired end sequencing libraries were prepared using the Illumina TruSeq Nano DNA Library Preparation kit. Metagenomic sequencing was performed using the Illumina HiSeq 2500 sequence platform at the Agricultural Research Council – Biotechnology Platform (ARC-BTP), Gauteng, South Africa, resulting in 250 nt paired–end reads (13,267,176 ± 3,534,751 reads per sample). Prior to assembly, the metagenomic reads were subject to adaptor removal and quality filtration using Trimmomatic (Bolger *et al*., 2014) with a minimum sliding window quality score of 20 and reads shorter than 100 bp were discarded. Following quality filtering, the level of coverage of each metagenome was assessed using Nonpareil, a statistical program where read redundancy is used to estimate coverage (Rodriguez and Konstantinidis, 2013). Prior to assembly, metagenomic reads were pooled and *de novo* assembly of quality trimmed reads into contiguous sequences (contigs) followed by scaffolding using metaSPAdes assembler version 3.9.0 (Nurk *et al*., 2017) with kmers list of 21, 33, 55, 77, 99, 127. The resulting assembly consisted of 1,007,176 scaffolds (> 500 bp) and an N50 and L50 of 1638 bp and 42230 bp, respectively. Reads were mapped to the scaffolds greater than 500 bp, bam files were filtered to retain mapping reads (samtools view –F 4), and the number of reads mapping to the scaffolds in the metagenomics assembly were counted using awk script. An average of 13,092,168 ± 3,561,160 reads per sample (> 500 bp) were mapped to scaffolds (i.e., 99 ± 1.6% of reads mapped to scaffolds) (Table C3).

Open reading frames (ORFs) on scaffolds were predicted using Prodigal (Hyatt *et al*., 2010) with the meta flag activated. The resulting predicted ORFs were annotated against KEGG (Kyoto encyclopaedia of genes and genomes) (Kanehisa *et al*., 2015) using DIAMOND (Buchfink *et al*., 2015). Genes involved in the nitrogen cycle were identified based on KEGG orthology (KO) numbers assigned to predict ORFs based on the KEGG nitrogen metabolism pathway (Table C4). The abundance (as reads per million kilobase (rpkm)) of genes was determined across all samples by dividing the number of reads mapping to scaffold containing the gene by the scaling factor (i.e., millions of reads per sample) and the length of the scaffold in kilobases. The quality filtered paired end reads were mapped to a database of 44 high quality complete and draft genomes of nitrifying organisms, AOA (n=11), AOB (n=19), comammox bacteria (n=4), NOB (n=5), and anammox bacteria (n=5) (Table C5). Reads were competitively mapped to reference genomes using bwa and properly paired reads

(samtools view –f 2) mapping to each reference were counted using awk script. The abundance of reference genomes in the samples was calculated by dividing the number of reads mapping to reference genomes by the scaling factor (i.e., millions of reads per sample) and the total length of the reference genome in kilobases. All raw sequence data have been deposited with links to BioProject accession number PRJNA524999 in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/).

2.4 Metagenome Assembled Genome (MAG) reconstruction

Assembled scaffolds (>2000 bp) from the co-assembly of all samples were used to generate metagenomic assembled genomes (MAGs) using CONCOCT (Alneberg et al., 2014). This resulted in the construction of 115 CONCOCT clusters. A total of 60 CONCOCT clusters with completeness greater than 50%, based on the occurrence of 36 single copy genes used by CONCOCT to estimate completeness, were selected for further examination/refinement. The completeness and redundancy of these 60 CONCOCT clusters was checked with CheckM (Parks et al., 2015) with 47 clusters selected for further analysis based on 75% completeness. Of these, three had redundancy estimates greater than 10% and were manually refined using Anvi'o (Eren et al., 2015). This resulted in 47 high quality Metagenome Assembled Genomes (MAGs) (>70% complete, less than 10% redundancy). Bins were functionally annotated and taxonomically classified using GhostKOALA (Kanehisa et al., 2016), where predicted ORF's were assigned KO numbers using KEGG's set of nonredundant KEGG genes. In addition, the abundance of all MAGs was calculated similar to that of the reference genomes (i.e., rpkm). Taxonomic annotation of the final MAGs was conducted using MiGA (Rodriguez and Konstantinidis, 2014). Taxonomic inference and characteristics of MAGs are detailed in Table C4. Genome-level inference of the 47 MAGs was then used to construct a phylogenomic tree using GToTree described by Lee (2019).

2.5 Marker gene based taxonomic and phylogenetic analysis

Small subunit (SSU) rRNA gene sequences were identified using a Hidden Markov Model (HMM) search using the Infernal package (Nawrocki et al., 2009) with domain specific covariance models and corrected as outlined previously (Brown et al., 2015). Detected SSU rRNA genes greater than 500 bp were classified using SILVA taxonomy and the relative abundance of each SSU rRNA gene was estimated by dividing the total coverage of the scaffold containing the SSU rRNA gene by the coverage of all scaffolds containing SSU

rRNA genes within each domain (i.e., bacteria, archaea, and eukaryota). Reference databases of ammonia moonooxygenase subunit A (*amoA*, KO: K10944) and nitrite oxidoreductase subunit A (*nxrA*, KO: K00370) genes were created using corresponding reference sequences obtained from NCBI GenBank Database with additional *nxrA* reference sequences obtained from Kitzinger *et al*. (2018). An alignment was created for each gene using MAFFT (version 7) online multiple alignment tool with the iterative refinement method L-INS-i (Katoh *et al*., 2002). Resulting alignments were examined and trimmed by removing all overhangs using BioEdit Sequence Alignment Editor v 7.2.6.1 (Hall, 2009), resulting in sequences of equal length. Aligned datasets were subjected to Maximum Likelihood analysis (Felsenstein, 1981) in MEGA7 (Molecular Evolutionary Genetic Analysis) (Kumar *et al*., 2016) using the best-fit substitution models as determined in MEGA7 model tests. For all Maximum Likelihood phylogenetic trees, branch support was estimated using non-parametric bootstrap analyses based on 1000 pseudoreplicates under the same model parameters and rooted with appropriate outgroups. Amino acid sequences of genes annotated as *amoA* and *nxrA* were then placed on the respective reference phylogenetic tree using pplacer (Matsen *et al*., 2010).

3. Results

3.1 Changes in ammonium, nitrite, and nitrate concentrations in the chloraminated section of the DWDS

Spatial changes in the concentrations of ammonium, nitrite and nitrate were observed as the chloraminated bulk water moved through the DWDS (Figure C1). Here, ammonium concentrations decreased as the bulk water moved away from the chloramination sites towards the end of the DWDS (i.e., average NH_4^+ concentrations decreased from 0.30 ± 0.12 mg/L following chloramination to 0.05 ± 0.09 mg/L at the end of the DWDS). The decrease in ammonium concentrations was associated with an increase in nitrite concentrations, which reached its highest concentrations at sites before RES2 (i.e., average NO_2^- concentrations increased from 0.005 ± 0.01 mg/L to $0.21 \pm .033$ mg/L) after which they decreased. Interestingly, both nitrite and nitrate levels increased in RES1, while increases in nitrate concentrations directly corresponded to decreases in nitrite concentrations after RES2. Nitrate concentrations peaked at locations after RES2 where ammonium and nitrite concentrations are the lowest (i.e., average NO_3^- concentrations increased from 0.18 ± 0.12 mg/L following chloramination to 0.40 ± 0.23 mg/L at the end of the DWDS).

A clear decrease in ammonium concentrations was observed after RES2 (Fig. 2B) (i.e., average ammonium concentrations of 0.30 ± 0.13 mg/L before RES2 and 0.16 ± 0.12 mg/L after RES2). In the case of RES1, increases in nitrite and nitrate concentrations were generally associated with concomitant decreases in ammonium concentrations in the first 8 months. However, these correlations between nitrite, nitrate and ammonium concentrations were not observed for the remaining months of the study period. Samples before RES2 had increased nitrite and nitrate concentrations and lower ammonium concentrations as compared to RES1. However, after RES2 nitrite levels dramatically decrease (i.e., average NO_2^- concentrations decreased from 0.17 ± 0.22 mg/L before RES2 to 0.09 ± 0.09 mg/L after RES2), which correlated with an increase in nitrate concentrations (i.e., average NO_3^- concentrations increased from 0.27 ± 0.13 mg/L before RES2 to 0.41 ± 0.22 mg/L after RES2).

Furthermore, ammonium, nitrite and nitrate concentrations demonstrated strong temporal trends in both reservoirs associated with changes in temperature and disinfectant residual concentrations with the highest ammonium concentrations in winter and spring months (~0.5 mg/l). Disinfection residual concentrations (i.e. total chlorine and monochloramine) were generally higher in winter and spring months (peaking in July 2016 at approximately 2.0 mg/L) and negative correlated with water temperature as shown previously (Potgieter *et al*., 2018) (Figure 2C). Nitrite and nitrate concentrations were highest in summer and autumn months (0.66 and 0.82 mg/L, respectively) where associated ammonium levels were low (0.12 mg/L). Typically, decrease in monochloramine and ammonium concentrations were associated with increased nitrite and/or nitrate concentrations at higher temperatures. It is important to note that the observed trends in ammonium, nitrite and nitrate concentrations may not be the same from year to year as increased levels of nitrification were observed for the first year (months 10-12) as opposed to second year (months 13-24) (Figure C2).

3.2 Changes in the microbial community composition between the two reservoirs

Based on SSU rRNAs identified in the metagenomic data, the majority of sequences identified were bacterial (i.e., the mean relative abundance (MRA) of bacterial SSU rRNA genes across all samples was 90.72 ± 7.23%), followed by unclassified SSU rRNA contigs (6.58 ± 5.29%) and eukaryota (2.70 ± 2.65%) with no archaeal SSU rRNA detected (Table C7). At the phylum level, *Proteobacteria* were the most abundant in both reservoirs (i.e.,

MRA of 81.60 ± 13.32% in RES1 and 82.98 ± 10.28% in RES2), followed by *Nitrospirae* with an MRA of 11.25 ± 13.48% in RES1, although the MRA of *Nitrospirae* decreased to 2.28 ± 3.83% in RES2 (Figure C3A). The decrease in abundance of both *Proteobacteria* and *Nitrospirae* in RES2 was associated with increases in both eukaryota (MRA: 2.05 ± 1.26% in RES1 to 4.64 ± 5.55% in RES2) and unclassified contigs (MRA: 3.55 ± 1.17% in RES1 to 7.58 ± 3.66% in RES2), of which 16.87% of unclassified contigs were less than 250 bp. The taxonomic classification of eukaryota 18S rRNA contigs is shown in Table C8.

Further classification of the proteobacterial classes based on SILVA taxonomy revealed that the abundance of *Alpha-* and *Gammaproteobacteria* (specifically order: *Betaproteobacteriales*) varied between RES1 and RES2. In RES1 the abundance of *Alphaproteobacteria* (MRA: 42.99 ± 13.97%) was marginally higher than *Betaproteobacteriales* (MRA: 38.05 ± 14.47%). However in RES2, the abundance of *Alphaproteobacteria* increased in dominance (MRA: 52.75 ± 15.36%) and *Betaproteobacteriales* decreased (MRA: 29.46 ± 17.14%) indicating a decrease in the abundance of *Betaproteobacteriales* as chloraminated water moves further down the DWDS. In RES2 there was an observed increase in the relative abundance of other *Gammaproteobacteria* from a MRA of 0.48 ± 0.37% in RES1 to 0.69 ± 0.35% in RES2 (Figure C3B). Furthermore, within *Betaproteobacteriales*, two *Nitrosomonas* SSU rRNAs (NODE_284 and NODE_310) were highly abundant with the mean relative abundance of 31.44 ± 15.18% in RES1 and 10.10 ± 10.07% in RES2, consequently making *Nitrosomonas* the most dominant genera identified in all samples. These results correlated with 16S rRNA gene profiling data previously reported by Potgieter *et al.* (2018).

Alpha diversity indices [richness (observed taxa), Shannon Diversity Index and Pielou's evenness] were calculated using the summary single function in mothur (Schloss et al., 2009) incorporating the parameters, iters=1000 and subsampling=475 (sample containing the least number of sequences). RES2 samples were more rich than RES1 samples (average number of taxa for RES1 was 49 ± 12, whereas for RES2 it was 69 ± 28) as well as slightly more diverse and even (i.e., Shannon Diversity Index: 2.91 ± 0.77 and Pielou's evenness: 0.69 ± 0.13) than RES1 (i.e., Shannon Diversity Index: 2.32± 0.44 and Pielou's evenness: 0.59 ± 0.10) (Figure C4). However, Kruskal-Wallis one-way analysis of variance and post-hoc Dunn's test revealed that differences in alpha diversity measures between the two reservoirs

were not significant (p > 0.05). Beta diversity measures (i.e. structure based: Bray-Curtis and membership based: Jaccard) between the two reservoirs for corresponding months revealed that the microbial communities between the two reservoirs were on average 60% dissimilar in community structure (Bray-Curtis: 0.59 ± 0.16) and 74% dissimilar in community membership (Jaccard: 0.74 ± 0.12). This indicated that the microbial community differs significantly in both community structure and membership between the two reservoirs (AMOVA, $F_{ST} \leq 2.12$, p < 0.05).

3.3 Dominant genes involved in nitrogen transforming reactions

Genes encoding for enzymes involved multiple nitrogen transforming reactions were observed across both reservoirs (Table C4). However, the coverage of some of these genes was low and as a result their contribution to overall nitrogen metabolism was thought to be limited in the chloraminated drinking water environment. Therefore, the dominant genes driving nitrogen transformation reactions were identified as those genes with a cumulative coverage of >100 reads per million kilobase (rpkm) in both reservoirs (Figure 3). These genes included *amoABC* and *hao* (ammonia oxidation), *nxrAB* (nitrite oxidation), *nasA* (assimilatory nitrate reduction), *nirBD* (assimilatory nitrite reduction), *nirK* (nitrite reduction, NO-forming) and *norBCDQ* genes (nitric oxide reduction). The cumulative coverage of these dominant genes for each sample from both reservoirs in shown in Figure 4.

3.3.1 Ammonia and nitrite oxidation

Genes encoding for the enzymes involved in nitrification were observed in all reservoir samples. Specifically, the genes responsible for ammonia oxidation were observed to have the highest coverage across all samples, out of all genes identified to be involved in the nitrogen cycle (Figure 4). In both reservoirs, *amoC*, *amoB* and *hao* genes had the highest total cumulative coverage in RES1 (i.e. 4625, 2922 and 2486 rpkm, respectively) and in RES2 (i.e. 4807, 2166 and 2170 rpkm, respectively). The average cumulative coverage of these dominant genes was higher across RES1 samples (i.e. *amoC*: 578 ± 287, *amoB*: 365 ± 206 rpkm and *hao*: 310 ± 150 rpkm) than across RES2 samples (i.e. *amoC*: 481 ± 458, *amoB*: 217 ± 171 rpkm and *hao*: 217 ± 172 rpkm). In addition, *amoA* genes were also observed to be highly abundant with a total cumulative coverage of 1143 rpkm in RES1 for all time points (average cumulative coverage of 143 ± 80 rpkm across all RES1 samples), which increased to 1672 rpkm in RES2 (with an average cumulative coverage of 167 ± 163 rpkm across all

RES2 samples) (Figure 3 and 4). Although, these genes showed high coverage, only a small number of genes were identified as *amoABC* and *hao*, indicating a low level of diversity among these genes (Figure C5). Here, the four individual contigs with the highest coverage across all samples (i.e. average coverage across all samples > 100 rpkm) contained the *hao* (NODE_10017), *amoC* (NODE_69075), *amoB* (NODE_58840) and *amoABC* (NODE_21034) genes. BLAST results revealed that the majority of *amoABC* and *hao* genes represented members of *Nitrosomonas* genus with a small minority of genes within each group were found to belong to *Nitrospira* species, indicating the potential presence of comammox bacteria in the bacterial community.

Genes encoding nitrite oxidoreductase (*nxrAB*), required for the oxidation of nitrite to nitrate, were also observed across all samples (Figure 4). As expected, BLAST results of *nxrA* and *nxrB* genes revealed that the majority of these genes represented *Nitrospira* species. Although the coverage of individual *nxrAB* genes was significantly lower than *amoABC* and *hao* genes, they still maintained a cumulative coverage of >100 rpkm in both reservoirs. *nxrA* genes showed a total cumulative coverage of 347 rpkm in RES1 (average cumulative coverage of 43 ± 45 rpkm across RES1 samples) and decreased to 183 rpkm in RES2 (with an average cumulative coverage of 18 ± 16 rpkm across RES2 samples). Similarly, *nxrB* genes showed a total cumulative rpkm of 335 rpkm in RES1 (average cumulative coverage of 42 ± 46 rpkm across RES1 samples) and decreased to 135 rpkm in RES2 (with an average cumulative coverage of 14 ± 17 rpkm across RES2 samples) (Figure 3 and 4). Specifically, the individual contig NODE_68 containing *nxrAB* genes was consistently present across all samples and contributed to the high cumulative coverage of *nxrAB* genes (with an average cumulative coverage of 39 ± 48 rpkm in RES1 and 9 ± 16 rpkm in RES2). The low coverage of *nxrAB* genes correlated with the small number of genes identified as *nxrAB* genes, indicating a low level of diversity within this function (Figure C5).

Although different in each reservoir, the temporal trends in the cumulative coverage revealed a contrasting relationship between *amoABC* and *nxrAB* genes (Figure C6A and C6B). In RES1, *nxrAB* genes showed increased coverage in March (2015 and 2016) relative to *amoA* genes. Conversely, months where *nxrAB* gene coverage decreased, the coverage of *amoABC* genes increased (April 2015 and 2016). This contrasting temporal trend was also observed in RES2.

3.3.2 Reduction of nitrate and nitrite

The nitrate formed through nitrification can be potentially reduced back to nitrite through nitrate reduction. Here, the cytoplasmic assimilatory nitrate reductase, *nasA* gene (encoding the catalytic subunit of the NADH-nitrate reductase) was identified as the most abundant nitrate reductase, indicating the capacity of certain members of the community to use nitrate as an alternative electron acceptor. Dissimilatory nitrate reductase genes were also identified, including respiratory membrane bound nitrate reductases (*narGHIJ*), although their cumulative coverage of was low (<100 cumulative rpkm in both reservoirs). The total cumulative coverage of all *nasA* genes identified was 455 rpkm in RES1 (average cumulative coverage of 57 ± 24 rpkm across RES1 samples) and increased to 701 rpkm in RES2 (with an average cumulative coverage of 70 ± 22 rpkm across RES2 samples) (Figure 3 and 4). BLAST analyses revealed that the majority of *nasA* genes represented members of *Alphaproteobacteria* (6% of *nasA* genes) of which 34% were identified as belonging to the order *Rhizobiales*. In addition, *Gammaproteobacteria* (order: *Betaproteobacteriales*) represented another 18% of *nasA* genes. Here, an increased number of genes within the metagenomic dataset were identified as *nasA*, converse to *amo* and *nxr* genes, indicating an increased level of diversity of microorganisms containing the cytoplasmic assimilatory nitrate reductase (Figure C5).

Genes encoding enzymes involved the assimilatory reduction of nitrite to ammonia (*nirAB* genes) were also observed to have high coverage across all reservoir samples (Figure 4). The ferredoxin-nitrite reductase (*nirA*) showed high abundance with a total cumulative coverage of 729 rpkm in RES1 (average cumulative coverage of 91 ± 41 rpkm across RES1 samples) and 629 rpkm in RES2 (with an average cumulative coverage of 69 ± 13 rpkm across RES2 samples) (Figure 3 and 4). BLAST analyses identified that 50% of *nirA* genes represented members of *Alphaproteobacteria* of which 36% were identified as *Rhizobiales* while 32% of *nirA* genes were identified as belonging to *Nitrospira* species. The nitrate reductase (NADH), large subunit (*nirB* gene) was also observed to have high coverage across both reservoirs (i.e. cumulative coverage >100 rpkm), although the abundance of *nirB* genes was less than that of the *nirA* genes with the total cumulative coverage of *nirB* genes being 249 rpkm in RES1 (average cumulative coverage of 32 ± 16 rpkm across RES1 samples) and increased to 438 rpkm in RES2 (with an average cumulative coverage of 44 ± 31 rpkm across RES2 samples) (Figure 4). BLAST results revealed that the majority of *nirB* genes represented

Gammaproteobacteria (order: *Betaproteobacteriales*) (69% of *nirB* genes) of which 28% were identified as *Nitrosomonadales* and 28% as *Burkholderiales*. In addition, another 21% of *nirB* genes were found to represent *Alphaproteobacteria* (order: *Rhizobiales)*. Although, both *nirAB* genes showed moderate to low coverage across both reservoirs, the number of genes identified as belonging to these functions was high, specifically *nirB* genes. This indicates that a diverse assemblage of bacteria had the genetic potential for assimilatory nitrite reduction (Figure C5).

However in RES1, temporal trends in the relative abundance of *nasA* and *nirA* genes showed a converse relationship (Figure C6C). Increased relative abundance of *nirA* genes (in February and March 2015 and March 2016) was associated with decreased relative abundance of *nasA* genes. However, this converse relationship was not observed between *nasA* and *nirB* in RES2. In RES2, the relative abundance of these genes (*nasA*, *nirA* and *nirB*) generally showed the same temporal trends across RES2 samples and the same converse relationship with ammonia concentrations, specifically between *nasA* and *nirA*. Increases in the relative abundance of both *nasA* and *nirA* genes was observed in April and May 2015 and March 2016 (Figure C6C). This suggests the potential for complete assimilatory nitrate reduction to ammonia (i.e., *nasA* and *nirAB*) specifically in RES2 as the relative abundance of *nasA* genes typically showed a converse relationship to ammonia concentrations.

Lastly, *nirK* genes encoding nitric oxide forming nitrite reductases (reducing nitrite to nitric oxide) were also observed to be highly abundant in both reservoirs (i.e. cumulative coverage >100 rpkm) (Figure 3). *nirK* genes were identified as the fourth most abundant gene (following *amoC*, *amoB* and *hao*, respectively) with total cumulative coverage of 2016 rpkm in RES1 (with an average cumulative coverage of 252 ± 92 rpkm across RES1 samples) and 2043 rpkm in RES2 (with an average cumulative coverage of 204 ± 114 rpkm across RES2 samples) (Figure 3 and 4). Furthermore, the cumulative coverage of *nirK* genes across all samples generally showed a converse relationship with *nirA* and *nirB* genes in both reservoirs (Figure C6C and C6D). In addition, a high number of genes were identified as *nirK,* indicating a high diversity within this function (Figure C5). Furthermore, BLAST results of individual *nirK* genes revealed that 36% of these genes belonged to species of *Nitrospira* and 17% to *Nitrosomonas* species. The remaining *nirK* genes were shown to represent a diverse

group of bacteria including both *Alpha-* and *Gammaproteobacteria* (order: *Betaproteobacteriales*).

3.3.3   Nitric oxide reduction

The nitric oxide formed through the reduction of nitrite can be further reduced to nitrous oxide by nitric oxide reductases (*norBCDEQ*). Genes encoding for nitric oxide reductases (*norBCDQ*) were identified to have high coverage in this dataset, across all samples (i.e. cumulative coverage >100 rpkm in both reservoirs) (Figure 3). These genes had higher total cumulative coverage in RES1 (i.e. *norB*: 1364 rpkm, *norQ*: 1122 rpkm, *norC*: 1060 rpkm and *norD*: 856 rpkm) than in RES2 (i.e. *norB*: 972 rpkm, *norQ*: 997 rpkm, *norC*: 427 rpkm and *norD*: 700 rpkm) (Figure 3 and 4). Interestingly, the temporal cumulative coverage of *norB* generally showed a converse relationship to the cumulative coverage of other *nor* genes, specifically *norC* and *norQ* (Figure C6E). The dynamics of *nor* genes may be associated with different members of the community therefore resulting in differences in their coverage. The diversity of *norBCQ* genes was generally higher than that of *norD* genes (Figure C5). BLAST analyses revealed that *norB* genes represented members of *Gammaproteobacteria* (order: *Betaproteobacteriales*) predominately the order *Nitrosomonadales* (33% of *norB* genes) and *Alphaproteobacteria,* predominately the family *Sphingomonaceae* (23% of *norB* genes). The majority of *norQ* genes were identified as *Nitrospira* (16% of *norQ* genes) and *Gammaproteobacteria* (order: *Betaproteobacteriales*) predominately the order *Nitrosomonadales* (40% of *norQ* genes), of which 14% was identified as *Nitrosomonas* species. Similarly, the majority of *norC* genes were identified as *Nitrospira* (44% of *norC* genes) and *Gammaproteobacteria* (order: *Betaproteobacteriales*) order *Nitrosomonadales* (36% of *norC* genes). Lastly, as with *norQ* and *norC* genes, 55% *norD* genes were found to represent members of *Gammaproteobacteria* (order: *Betaproteobacteriales*) order *Nitrosomonadales* (of which 21% was identified as *Nitrosomonas* species) and 13% as members of the genus *Nitrospira*.

3.3.4   Other processes involved in nitrogen metabolism

Briefly, genes involved in other processes in nitrogen metabolism, with very low coverage across both reservoirs (< 100 cumulative rpkm) also were identified (Table C4). These genes included nitrous oxide reductase (*nosZ*), nitrogenases involved in nitrogen fixation (*nif* genes) as well as other genes involved in nitrite and nitrate transport (i.e. *nrtABC* genes encoding the

enzymes periplasmic substrate-binding protein, permease protein and an ATP-binding protein, respectively). The presence of *nrtABC* often co-occurred with presence of *nasA*, suggesting these nitrate transport enzymes are associated with assimilatory nitrate reductases. In addition, the gene encoding nitronate monooxygenase (*ncd*2) was also identified suggesting the potential to convert nitronate to nitrite and potentially provide an alternative source of nitrite.

3.4 Nitrifier diversity based on phylogenetic inference of *amoA* and *nxrA* genes.

Given the high abundance of *Nitrosomonas* and *Nitrospira,* the phylogenetic diversity of *amoA* and *nxrA* genes (marker genes for ammonia and nitrite oxidation, respectively) was investigated. Consistent with BLAST results, the phylogenetic placement of *amoA* genes revealed that nearly all *amoA* genes clustered with *Nitrosomonas* species indicating that the majority of *amoA* genes are associated with strict ammonia oxidisers (Figure 5A). Of these genes, three grouped closely with *Nitrosomonas oligotropha,* one with *Nitrosomonas* sp. AL212 and one with *Nitrosomonas* sp. Is79A3. The clustering of the remaining contigs remained within the *Nitrosomonas* but did not group closely with specific species (Figure 5A). Interestingly, a single *amoA* contig (NODE_6011) grouped within the *Nitrospira* comammox cluster. Furthermore, BLAST results confirmed that *amoA* contig (NODE_6011) represented a *Nitrospira* species, indicating the potential presence of a comammox bacteria within the microbial community.

Also consistent with BLAST results, phylogenetic placement of the *nxrA* genes revealed that the majority *of nxrA* genes belonged to the widely distributed lineage II of the genus *Nitrospira* grouping with the canonical nitrite oxidisers, *N. lenta* and *N. moscoviensis* and *Ca.* Nitrospira sp. ST-bin5 described by Wang *et al*. (2017) (Figure 5B). In addition, some *nxrA* genes grouped closely with known comammox *Nitrospira,* including *Ca*. Nitrospira nitrosa, *Ca*. Nitrospira inopinata and *Ca*. Nitrospira nitrificans. However, identification and taxonomic resolution of comammox bacteria based on *nxrA* gene phylogeny is not possible. Lastly, a single *nxrA* gene grouped closely with the anammox *Ca.* Scalindua brodae although no other evidence of anammox bacteria was found in the metagenomic data. Furthermore, no *nxrA* genes were found to be associated with *Nitrobacter* species (Figure 5B).

3.5 The nitrogen metabolic potential of dominant Metagenome Assembled Genomes (MAGs)

Following metagenomic binning, the 47 high quality MAGs constructed included 25 *Alphaproteobacteria*, 15 *Betaproteobacteriales*, 2 *Nitrospirae* as well as another 5 MAGs identified as *Bacteroidetes*, *Gammaproteobacteria*, *Gemmatimonadetes* and 2 *Planctomycetes,* respectively (Figure 6)*.* MAGs were genetically compared to their closest related reference genomes based on average amino acid identity (AAI). The complete description of the 47 MAGs and their coverage across the two reservoirs are shown in Table C4 and Figure C3. This analysis suggested that the majority of MAGs constructed represent genomes that are not yet represented in public database. However, the classification of these MAGs identified members of several abundant taxa identified in this chloraminated DWDS described by Potgieter *et al*. (2018). The 47 constructed draft genomes were screened for key functional genes involved in KEGG nitrogen metabolism pathway and the genetic potential for nitrogen transforming reactions within these MAGs was found to be diverse (Figure 6).

Of the 47 MAGs, 5 MAGs were observed to dominate the community with a cumulative rpkm of >100 rpkm across both reservoirs (Figure 7). These MAGs were identified as two *Nitrosomonas*-like MAGs (C58 and C107), a *Rhizobiales*-like MAG (C103.2), a *Sphingomonas*-like MAG (C70) and a *Nitrospira*-like MAG (C51). Furthermore, the relative abundance of these top 5 most abundant SSU rRNAs in the metagenomic dataset were correlated with the rpkm based abundance of these MAGs showing the same trends across both reservoirs (Figure C8) and the contigs containing these SSU rRNAs were observed in their corresponding MAGs.

3.5.1   *Nitrosomonas*-like MAGs (C58 and C107)

Contigs containing the *Nitrosomonas* SSU rRNAs were found to be associated with the two dominant *Nitrosomonas*-like MAGs [i.e. contig NODE_284 with the *Nitrosomonas*-like MAG (C58) and contig NODE_310 with the *Nitrosomonas*-like MAG (C107)]. Phylogenetic analysis, of these *Nitrosomonas* spp. SSU rRNA contigs revealed that they both clustered with *Nitrosomonas oligotropha,* with good bootstrap support (Figure C9). The classification of the *Nitrosomonas* SSU rRNAs as *Nitrosomonas oligotropha* correlates with the high coverage observed of the *Nitrosomonas oligotropha* reference genome across all samples (i.e., average coverage of 47 ± 23 rpkm in RES1 and 29 ± 24 rpkm in RES2) (Figure C11).

Thus, there may be two separate populations of *Nitrosomonas* responsible for the first step in nitrification.

Of the dominant MAGs, *Nitrosomonas*-like MAG (C107) exhibited the highest abundance in the metagenomic dataset. Although the two *Nitrosomonas*-like MAGs were potentially identified as the same species (based on SSU rRNAs), differences in the average coverage of these MAGs across the two reservoirs indicated that they may exhibit competitive dynamics (Figure C7 and C8). Here, the coverage of *Nitrosomonas*-like MAG (C58) increased from RES1 (7 ± 10 rpkm) to RES2 (22 ± 27 rpkm), whereas *Nitrosomonas*-like MAG (C107) decreased in coverage from RES1 (43 ± 26 rpkm) to RES2 (12 ± 6 rpkm) (Figure 7). Spearman correlations revealed that there was a moderate negative correlation between these two MAGs in RES1, however this was not statistically significant (Spearman correlation: -0.40, $p = 0.3268$).

Within both the *Nitrosomonas*-like MAGs, genes required for ammonia oxidation, i.e., ammonia monooxygenase (*amoABC*) and hydroxylamine oxidoreductase (*hao*) were observed. More specifically, the dominant genes in the metagenomic dataset, *hao* (NODE_10017) and *amoABC* (NODE_21034) could be linked to the *Nitrosomonas*-like MAG (C107). Furthermore, the *amoA* contig (NODE_21034) was observed to group closely with *N. oligotropha* (Figure 5A). However, an *amoA* gene was not recovered from *Nitrosomonas*-like MAG (C58) and this may be due to its overall lower abundance. In addition, both *Nitrosomonas*-like MAGs also contained *nirK* genes (nitrite reductase, nitric oxide forming) and *nor* genes (nitric oxide reductases) (Figure 6). More specifically, *Nitrosomonas*-like MAGs (C107) contained multiple *nor* genes including *norBCDQ*, whereas *Nitrosomonas*-like MAGs (C58) only contained *norQ*, suggesting that these dominant MAGs may play a role in regulating the concentrations of nitric oxide.

Spearman correlations between *Nitrosomonas*-like MAG (C107) and other dominant MAGs revealed strong correlations specifically in RES2. Strong positive correlations were observed with *Sphingomonas*-like MAG (C70) (Spearman correlation: 0.76, $p < 0.05$) and *Rhizobiales*-like MAG (C103.2) (Spearman correlation: 0.72, $p < 0.05$). In addition, *Nitrosomonas*-like MAG (C58) showed a strong positive correlation with *Rhizobiales*-like MAG (C103.2) (Spearman correlation: 0.81, $p < 0.05$) (Figure C8).

3.5.2   *Nitrospira*-like MAG (C51)

The *Nitrospira*-like MAG (C51) was dominant across all samples showing the 5th highest average coverage (10 ± 15 rpkm). The coverage of *Nitrospira*-like MAG (C51) was higher across RES1 samples (17 ± 20 rpkm) but decreased in RES2 (4 ± 69 rpkm) (Figure 7). This same trend was observed with the *Nitrospira* SSU rRNA contig (NODE_2) as the contig was identified in *Nitrospira*-like MAG (C51). Following phylogenetic analysis of *Nitrospira* SSU rRNAs, it was observed that *Nitrospira* spp. SSU rRNA NODE_2 grouped closely with *Nitrospira lenta* (lineage II) (Figure C10). The classification of the dominant *Nitrospira* SSU rRNA as *Nitrospira lenta* correlated with the observation of high coverage of the *Nitrospira lenta* reference genome across all samples with an average coverage of 14 ± 15 rpkm in RES1 and 4 ± 5 rpkm in RES2 (Figure C11). Although the SSU rRNA (NODE_2) was identified as *Nitrospira lenta* and this *Nitrospira*-like MAG (C51) does not contain *amoABC* or *hao* genes, further investigation into the potential that this MAG was a comammox revealed that it does contain genes for cytochrome *c* biogenesis, *ccmEFGH* genes, thought to be specific to comammox *Nitrospira*. However, BLAST results revealed that these *ccmEFGH* genes represented *Nitrospira lenta* species and these genes were present in *Nitrospira lenta* reference genome, which was consistent with the taxonomy associated with the SSU rRNA gene in the *Nitrospira*-like MAG (C51).

In addition to the highly abundant nitrite oxidoreductase genes (*nxrAB*) (NODE_68), this dominant *Nitrospira*-like MAG (C51) had the potential for other nitrogen transforming reaction as it contained genes involved in assimilatory nitrate and nitrite reduction (*nasA* and *nirA*, respectively) as well as the nitric oxide forming nitrite reductase (*nirK*) (Figure 6). This suggests that this MAG has the potential for complete assimilatory nitrate reduction to ammonia (i.e., *nasA* and *nirA*).

Correlations with other dominant MAGs revealed that within RES2, *Nitrospira*-like MAG (C51) had a strong negative correlation with *Nitrosomonas*-like MAG (C107) (Spearman correlation: -0.98, $p < 0.001$). Interestingly, *Nitrospira*-like MAG (C51) showed a strong negative correlation to another *Nitrospira*-like MAG (C56) in RES1 (Spearman correlation: -0.79, $p < 0.05$), although this other *Nitrospira*-like MAG (C56) was present at very low coverage across all samples. Additional correlation analyses revealed strong negative correlations between *Nitrospira*-like MAG (C51) and *Sphingomonas*-like MAG (C70) and

Rhizobiales-like MAG (C103.2) (C70: Spearman correlation: -0.73, p < 0.05 and C103.2: Spearman correlation: -0.64, p < 0.05) (Figure C8).

3.5.3   *Rhizobiales*-like MAG (C103.2)

The *Rhizobiales*-like MAG (C103.2) showed high coverage across both RES1 (25 ± 12 rpkm) and RES2 (20 ± 10 rpkm) constituting the second most abundant MAG across all samples (Figure 7). The genetic potential of this MAG in terms of the nitrogen cycle included assimilatory nitrate and nitrite reduction (*nasA* and *nirAB*, respectively) also indicating the potential of this MAG for complete assimilatory nitrate reduction to ammonia (i.e., *nasA* and *nirA*) (Figure 6). The poor taxonomic resolution of this MAG could not be improved based on the contig containing the SSU rRNA (NODE_1241) as it was also only identified as *Rhizobiales*. However, BLAST results of the *nasA* and *nirAB* genes revealed a close relation to *Proteobacteria* bacterium ST_bin 15 reported by Wang *et al*. (2017). Furthermore, assessment of additional correlations between *Rhizobiales*-like MAG (C103.2) and other dominant MAGs showed a strong positive correlation with *Sphingomonas*-like MAG (C70) (Spearman correlation: 0.76, p < 0.05) (Figure C8).

3.5.4   *Sphingomonas*-like MAG (C70)

Lastly, *Sphingomonas*-like MAG (C70) showed consistent coverage across both reservoirs (i.e. RES1: 11 ± 18 rpkm and RES2: 12 ± 18 rpkm) constituting the fourth most abundant MAG within the community (Figure 7). However, the genetic potential for the involvement of this MAG in nitrogen transformations was limited as it was observed to only contain the nitric oxide reductase *norB* gene (Figure 6). This corresponds to the increased number of *norB* genes identified (through BLAST analysis) as members of the family *Sphingomonaceae*. Further taxonomic classification of both the SSU rRNA contig and *norB* gene identified in this MAG, confirmed its taxonomy as *Sphingomonas*. The dominance of this MAG within the community suggests that it may be involved in other important metabolic capabilities within the microbial community that extend beyond the nitrogen cycle, such as biofilm formation.

4. Discussion

This study represents an in-depth investigation into the microbial nitrogen metabolism in chloraminated drinking water reservoirs. In doing so, we provide insight into the genetic

potential of a chloraminated drinking water microbial community to transform nitrogen species and identify the dominant members responsible for specific pathways within the nitrogen cycle. Previous studies have generally concentrated on the relationship between chloramination and nitrification (Regan *et al.,* 2003; Zhang *et al.,* 2009; Sawade *et al.*, 2016; Moradi *et al.*, 2017). However, this study comprehensively explores nitrogen metabolism, where changes in the microbial community composition as well as the genetic potential in terms of nitrogen metabolism is revealed in two linked chloraminated reservoirs.

4.1 Varying stages of nitrification between the two reservoirs

The chemical monitoring data indicated differing patterns in the concentrations of nitrogen species and spatial changes in disinfectant residuals in the two reservoirs, suggesting different nitrification regimes in the two reservoirs. Varying stages of nitrification in different sections of a chloraminated DWDS was also previously observed by Shaw *et al.* (2015). Nitrification was observed, specifically in RES2, as decreases in ammonium concentrations were associated with concomitant increases in nitrite and nitrate concentrations. An increase in the concentrations of both nitrite and nitrate were observed following RES1, although their concentrations were low and did not consistently correlate to changes in ammonium concentrations. While nitrification occurred in both reservoirs, it did not occur to a significant extent in RES1 for all sampled time points. This may be a consequence of increased monochloramine concentrations, as months with increased monochloramine concentrations showed reduced nitrification rates. In addition, ammonium concentrations were consistently higher in RES1 compared to RES2, as expected as these samples were closer to the site of chloramination. Conversely, following RES2, samples typically had elevated nitrate concentrations with reduced nitrite and ammonium concentrations, suggesting that with increasing distance from the site of chloramination, samples nearing the end of the DWDS undergo complete nitrification. Here it is more likely that the depletion of nitrite may be the result of tight coupling of canonical ammonia and nitrite oxidation (Shaw *et al.*, 2015; Daims *et al.*, 2016).

4.2 Changes in bacterial community composition between the two reservoirs

The majority of annotated proteins were bacterial and the taxonomic profiles based on SSU rRNA genes were in agreement with previous descriptions of the microbial community within chloraminated DWDS, in which *Proteobacteria* and *Nitrospira* (to a lesser extent)

were the dominant phyla (Bautista-de los Santos *et al*., 2016; Gülay *et al*., 2016: Potgieter *et al.,* 2018). The change in dominance between *Alphaproteobacteria* and *Betaproteobacteriales* in RES1 and RES2 was also observed by Potgieter *et al*., 2018, where *Betaproteobacteriales* were more abundant closer to the site of chloramination and an increase in *Alphaproteobacteria* correlated to decreased disinfectant residual concentrations in RES2. The variation in dominance between *Alphaproteobacteria* and *Betaproteobacteria* (now classified as the order *Betaproteobacteriales* in the phylum *Gammaproteobacteria*) has been well documented, where their dominance varies depending on multiple factors including disinfectant residual concentrations (Gomez-Alvarez *et al*., 2012; Wang *et al*., 2014) and seasonal trends (McCoy and VanBriesen, 2012, 2014; Pinto *et al*., 2014; Prest *et al*., 2016; Zlatanovic *et al*., 2017). Therefore, in this study, the increased retention time in the reservoirs may have an effect on the microbial community composition and function in terms of nitrogen transformation.

However, due to the lack of quantitative and viability assays it is unclear what proportion of the community data is from viable or active cells. Although there is a lack of absolute abundance in this study, it does not detract from the observed genetic potential of the microbial community to transform nitrogen. Furthermore, the coverage of the highly abundant MAGs (specifically, *Nitrosomonas, Sphingomonas* and *Nitrospira*) corresponds to that found by Sakcham *et al*. (2019) where after removal of extracellular DNA (eDNA), *Nitrosomonas* was present at greater coverage than *Nitrospira.* Furthermore, at a location with higher nitrite concentrations and after eDNA removal, *Sphingomonas* showed an increase in relative abundance, which correlated to higher abundances of *Nitrosomonas* at this location. This positive correlation between *Sphingomonas* and *Nitrosomonas* was also observed in this study. Here, *Sphingomonas* may act as a potential indicator for the onset of nitrification in chloraminated systems (Sakcham *et al*., 2019).

4.3 Nitrification driven by co-occurring *Nitrosomonas* and *Nitrospira* species

The engineered oligotrophic environment of drinking water systems typically have limited substrate for nitrifiers. However, in chloraminated drinking water systems, nitrogen (in the form of ammonia) is available either as excess ammonia or through disinfectant decay. This addition of ammonia consequently promotes the growth of nitrifying bacteria and drives nitrification (Regan, 2003; Wang *et al*., 2017). The phylogenetic distribution of most

commonly occurring ammonia-oxidising bacteria and archaea has been well described with AOB falling within *Gammaproteobacteria* (order: *Betaproteobacteriales*), the AOA falling specifically within the *Thaumarchaea* (Pester *et al*., 2011). However, in this study *Thaumarchaeota* was not detected. AOA are thought to play a more significant role in nitrification in environments that are low in dissolved oxygen or low ammonia concentrations (Schleper *et al*., 2010; Pester *et al*., 2011). In contrast, *Betaproteobacteriales* AOB, specifically *Nitrosomonas* were present in high abundances, indicating that *Nitrosomonas* is a dominant member of the microbial community and plays a significant role in the fate of nitrogen in this chloraminated system. The finding of AOB dominance over AOA is consistent with other chloraminated DWDS studies (de Vet *et al*., 2011; Wang *et al*., 2014).

Nitrification in chloraminated drinking water systems has been well characterised with *Nitrosomonas* and *Nitrospira* species often identified as the major ammonia- and nitrite-oxidisers, respectively (Regan, 2003; Hoefel *et al*., 2005; Zhang *et al*., 2009). In this study, the dominance of *Nitrosomonas* and *Nitrospira* species was confirmed as these species showed: (i) high abundance of their respective SSU rRNAs, (ii) high coverage of nitrifying genes associated with *Nitrosomonas* (*amoABC* and *hao* genes) and *Nitrospira* (*nxrAB* genes), (iii) high coverage of their respective *Nitrosomonas* and *Nitrospira* reference genomes and lastly (iv) high coverage of constructed *Nitrosomonas*-like and *Nitrospira*-like MAGs, across both reservoirs. The dominance of *Nitrosomonas* species in this study corresponded to the high abundance of *Nitrosomonas* observed by Potgieter *et al*. (2018), where they observed *Nitrosomonas*-like OTUs to be the most abundant OTUs, constituting 18.3% of the total sequences. Furthermore, *Nitrospira* was identified as the second most abundant phyla in this study, which correlated to the observed increase of *Nitrospira* in the reservoir samples observed by Potgieter *et al*. (2018).

Here, the chemical monitoring data alone may be limited in conclusively confirming the occurrence of nitrification however, the increased concentrations of nitrate together with the high abundance of *Nitrosomonas* and *Nitrospira* strongly suggests that nitrification occurs within both chloraminated reservoirs. Furthermore, Duff *et al*. (2017) reported that *amoA* gene abundances showed significant correlations to the Potential Nitrification Rate (PNR) from AOB in marine inertial bays. Here the high abundance of genes involved in ammonia-oxidation correlates to the observed increase in nitrate concentrations.

The phylogeny of the *amoA* gene confirmed that the majority of AOB were identified as strict ammonia oxidisers (specifically *Nitrosomonas*) with the presence of one potential comammox *amoA* distinct from the canonical AOBs. Some AOB representatives, such as *Nitrosomonas oligotropha,* may have a higher affinity for ammonia (Regan, 2003). Given the low concentrations of ammonia in these systems, the high affinity of *Nitrosomonas oligotropha* may allow it to outcompete other *Nitrosomonas* species (Regan, 2003). This is likely to be the case in RES2 where ammonium concentrations are lower than in RES1. Here, where the *Nitrosomonas*-like MAG (C107) decreases in coverage, the other *Nitrosomonas*-like MAG (C58) increases, indicating that C58 may have a higher affinity for ammonia when ammonia concentrations are low and is therefore able to outcompete other *Nitrosomonas*-like population.

Similarly, the phylogeny of *nxrA* genes showed that the majority of *nxrA* genes group within *Nitrospira* lineage II, which is the most widespread and diverse of the lineages, including both canonical NOBs as well as all currently known comammox bacteria (Koch *et al*., 2015; Daims *et al*., 2016; Daims and Wagner, 2018). The majority of *nxrA* genes from this study grouped closely with *Ca.* Nitrospira sp. ST-bin5 described by Wang *et al.* (2017) and the grouping of *Ca.* Nitrospira sp. ST-bin5 with *Nitrospira moscoviensis* was observed in both this study and the study by Wang *et al.* (2017). Furthermore, in their study they describe that *Ca.* Nitrospira sp. ST-bin5 does not contain the genes responsible for ammonia oxidation (*amo* and *hao*) and is therefore not a true comammox bacteria. However, it is not possible to identify the presence of comammox bacteria based on *nxrA* phylogeny.

It is now known that comammox and canonical AOB both use ammonia as substrate and therefore may co-exist in niches with low ammonia concentrations such as drinking water (Wang *et al*., 2017). Therefore, investigations into the presence of comammox revealed potential comammox based on *amoA* phylogeny and the identification of cytochrome c biogenesis genes (*ccmEFGH*) in both *Nitrospira*-like MAGs (Palomo *et al*., 2018). However, *ccmEFGH* genes were found to belong to *Nitrospira lenta* species, which was consistent with the taxonomy of the SSU rRNA within the *Nitrospira*-like MAG (C51). In addition, cyanate hydratase or cyanase (*cynS*) and assimilatory nitrite reductase (*nirA*) genes were observed in both *Nitrospira*-like MAGs. These genes are typically absent in comammox and only

detected in canonical NOB *Nitrospira,* which further confirms that the *Nitrospira*-like MAG (C51) is most likely related to *Nitrospira lenta.*

Furthermore, the dominance of one *Nitrospira*-like MAG (C51) was observed over another *Nitrospira*-like MAG (C56) indicating its ability to outcompete other *Nitrospira* members similar to what has been demonstrated in activated sludge systems (Maixner *et al.*, 2006; Ushiki *et al.*, 2017). Competition and separation in ecological niches between *Nitrospira* may be caused by physiological properties such as the affinity for nitrite and other substrates, formate utilisation and the relationship with AOB (Ushiki *et al.*, 2017). Typically, nitrification occurs in a modular fashion performed by a complex network of specialised microorganisms. This modularity results in the cooperative and competitive interactions (Stein and Klotz, 2016; Kuypers *et al.*, 2018). Nitrification generally occurs through a cooperative interaction between ammonia and nitrite oxidisers. Competitive interactions may exist within the ammonia-oxidising and nitrite-oxidising groups, which potentially compete for substrates ammonia and nitrite, respectively (Kuypers *et al.*, 2018). Here, the dominant *Nitrospira*-like MAG (C51) showed a positive relationship with the *Nitrosomonas*-like MAG (C107), decreasing in coverage in RES2 where disinfectant residuals and consequently ammonium concentrations were lower. This relationship may involve the tight coupling of canonical ammonia and nitrite oxidation, thereby resulting in the low concentrations of nitrite observed in both reservoirs (Shaw *et al.*, 2015; Daims *et al.*, 2016).

4.4 Genetic potential of the microbial community for nitrogen metabolism

Despite the limitations in available drinking water metagenomes, the results in this study provide insight into the diversity of microorganisms involved in nitrogen transformation in chloraminated drinking water. As a consequence of nitrification, concentrations of nitrate increase, thereby providing an alternative form of biologically available nitrogen. The increased availability of nitrate potentially promotes the growth of a highly diverse assemblage of microorganisms. There is an astounding diversity of microorganisms that transform nitrogen, where each microorganism has specific physiological requirements for optimal growth. Nitrogen transformations in the environment are typically carried out by microbial communities that recycle nitrogen more efficiently than a single microorganism. These microbial communities retain their nitrogen transforming capabilities even when the community composition is altered with changes in the environment. These microbial nitrogen

transforming reactions form complex networks in both the natural and engineered environments (Kuypers *et al*., 2018).

Where nitrification is driven by a select few microorganisms (i.e., *Nitrosomonas* and *Nitrospira*), the reduction of nitrate, nitrite and nitric oxide was likely performed by a diverse assembly of bacteria (predominately *Alphaproteobacteria*, *Gammaproteobacteria,* order *Betaproteobacteriales* and *Nitrospira*), each with their own discrete physiological requirements for optimal growth (Kuypers *et al*., 2018). The dominance of *amoABC* and *hao* (ammonia oxidation), *nxrAB* (nitrite oxidation), *nasA* (assimilatory nitrate reduction), *nirBD* (nitrite reductase), *nirK* (nitrite reductase, NO-forming) and *norBCDQ* (nitric oxide reduction) genes suggests that ammonia is oxidised and the resulting in the formation of nitrite and nitrate. The resulting oxidised compounds are most likely either (i) assimilatoraly reduced to nitric oxide and ultimately nitrous oxide potentially triggering biofilm formation or (ii) when ammonia concentrations are low, nitrite and nitrate are fixed and converted to ammonia for assimilation (Figure 8).

Bacterial nitrate reduction has been shown to be a multifaceted process, performed by three distinct classes of nitrate reducing systems differentiated by their cellular location, regulation, structure, chemical properties and gene organisation (Moreno-Vivian *et al*., 1999). In this study, all three systems (i.e., Nas, Nar and Nap) were observed within the diverse group of recovered draft genomes. The diversity observed within nitrate reductases in this study correlated with that observed in other studies where nitrate reductases were found to phylogenetically widespread (Richardson *et al*., 2001; Philippot, 2005; Smith *et al*., 2007). In many cases in this study, a single draft genome contained genes for both assimilatory and dissimilatory nitrate reduction, indicating that these pathways may be interconnected where the enzymes may play different roles under different metabolic conditions (Moreno-Vivian *et al*., 1999). The interconnection of assimilatory, respiratory and dissimilatory nitrate reduction may facilitate rapid adaptation to changes in nitrogen and/or oxygen conditions thereby increasing the metabolic potential for survival in the chloraminated drinking water environment (Moreno-Vivian *et al*., 1999).

In this study, the potential for assimilatory nitrate reduction was the dominant nitrate reducing pathway as *nasA* genes were highly abundant and were the only nitrate reductase

genes observed in the dominant MAGs, i.e. the *Nitrospira*-like MAG (C51) and the *Rhizobiales*-like MAG (C103.2). Assimilatory nitrate reductases are typically cytoplasmic and enables the utilisation of environmental nitrate as a nitrogen source. This enzyme is generally induced by nitrate but inhibited by ammonium, however it is not affected by oxygen (Moreno-Vivian *et al*., 1999). This converse relationship was observed between the abundance of *nasA* genes and ammonia concentrations. The high abundance of genes associated with assimilatory nitrate reduction in this study correlated with reduced concentrations of ammonium and increased concentrations of nitrate, which is a result of nitrification. In the oligotrophic environment of DWDS, where nutrients are limited, utilization of nitrate as nitrogen source for biomass synthesis may be an important mechanism for survival (Rivett *et al*., 2008).

The metabolic fate of nitrite formed through the reduction of nitrate may involve subsequent reduction to ammonia (through assimilatory nitrite reduction) or to nitric oxide (Moreno-Vivian *et al*., 1999; Smith *et al*., 2007; Daims *et al.*, 2016). Both *nirA* and *nirB* genes encoding nitrite reductases were observed to have high coverage across both reservoirs. The positive correlation with *nasA* and *nirA* confirms the potential for complete assimilatory nitrate reduction to ammonia, specifically in RES2. However in RES1, a converse relationship was observed between *nasA* and *nirA* genes, suggesting that *nirB* may potentially may play a larger role in the reduction of nitrite here as *nirB* genes showed similar trends to *nasA* genes in RES1, although, at a lower coverage than *nirA*. The NirB nitrite reductase uses NADH as an electron donor to reduce nitrite in the cytoplasm and typically, high nitrite concentrations are needed for *nirB,* which is consistent with the proposed role of the NirB enzyme in detoxification of nitrite. This may be the case following RES1, where nitrite concentrations were generally higher as compared to RES2.

Genes encoding dissimilatory nitrate reduction and denitrification were identified in this study, however their coverage across both reservoirs was generally very low. Both dissimilatory nitrate reduction to ammonia (DNRA) and denitrification compete for the nitrate and nitrite as an electron acceptor (Smith *et al*., 2016). It has been shown that DNRA may be favoured when nitrate concentrations are low and organic electron donor availability is high, whereas denitrification outcompetes DNRA when nitrate concentrations are high and carbon supplies are limiting (Rivett *et al*., 2008; Smith *et al*., 2016). However, denitrification

typically occurs in environments where available oxygen is limited and nitrate is used in respiration. Therefore, denitrification may not play a role in an aerobic drinking water environment.

Alternatively, nitrite may be reduced to nitric oxide (NO) via *nirK* (nitrite reductase, NO-forming), which showed high coverage across both reservoirs and was observed in *Nitrosomonas*-like and *Nitrospira*-like MAGs. At low, non-toxic concentrations, NO can potentially elicit other cellular responses other than denitrification (Arora *et al*., 2015). Although it is difficult to separate NO signalling responses from detoxification and denitrification, in *Nitrosomonas europaea* it was observed that low concentrations of NO caused biofilm dispersal, whereas high concentrations of NO caused increased biofilm formation as a defence mechanism (Arora *et al*., 2015). In this study, generally a converse relationship was observed between assimilatory nitrite reduction (*nirAB*) and NO forming nitrite reduction (*nirK*) suggesting that potentially when ammonia concentrations are increased, the need for nitrite to be assimilated to ammonia is reduced and nitrite may be preferentially reduced to nitric oxide. Furthermore, *nor* genes were observed to have high coverage across reservoir samples and were present in both *Nitrosomonas*-like MAGs and the *Sphingomonas*-like MAG. *Nor* genes are responsible for regulating the concentration of NO by the reduction of NO to nitrous oxide (N$_2$O). As denitrification may not be an important process in this system and the observed high coverage of *nirK* and *nor* genes, suggests that production of NO may act as an important molecule in regulation biofilm formation.

4.4.1   Other nitrogen transforming processes

The metabolic versatility of nitrogen transforming processes included nitrogen fixation. Although *nif* genes were observed at very low coverages, a full cascade of *nif* genes were identified in a *Sideroxydans*-like MAG, which is well described as an iron-oxidising bacteria. *Sideroxydans*, has previously been described as commonly occurring OTU in disinfectant free drinking water (Bautista-de los Santos *et al*., 2016) however, the presence of these iron-oxidisers in DWDS are typically, linked to microbial mediated corrosion of the steel and iron pipes (Emerson and De Vet, 2015). Further investigations of the interactions between iron-oxidation and nitrogen fixation may be needed for the optimisation of the design and operations of chloraminated DWDS.

5. Conclusion

The transformation of nitrogen has been well characterised in many environments including the open oceans and ocean sediments, wastewater and agricultural soils. However, to our knowledge this information is limited when considering chloraminated DWDSs, where nitrification is a major concern. In this study, *Nitrosomonas* and *Nitrospira* were identified as the main drivers in nitrification and together with the water chemistry data (i.e., changes in ammonia, nitrite and nitrate concentrations), this study improves our understanding nitrogen cycling potential in chloraminated drinking water. Furthermore, the genes and bacteria responsible for the metabolic fate of nitrate and other nitrogen transforming reactions were identified and were shown to be highly diverse. Here, the addition of ammonia through chloramination promotes/supports the growth of nitrifying bacteria and in this DWDS, the nitrate formed through nitrification may ultimately be reduced via assimilatory processes either to ammonia when ammonia concentrations are low or to nitric oxide for potential regulation of biofilm formation. This study therefore provides insight into the genetic network behind microbially mediated nitrogen metabolism in chloraminated drinking water.

6. References

1. Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., Lahti, L., Loman, N. J., Andersson, A. F. and Quince C. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods*. 11:1144-1146. doi:10.1038/nmeth.3103.

2. Arora, D. P., Hossain, S., Xu, Y. and Boon, E. M. (2015). Nitric oxide regulation of bacterial biofilms. *Biochemistry*. 54(24): 3717-3728.

3. 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.

4. Belser, L. W. (1976). Population ecology of nitrifying bacteria. *Annual Review of Microbiology*. 33: 309-333.

5. Bolger, A. M., Lohse, M. and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30(15): 2114-2120.

6. Brown, C. T., Hug, L. A., Thomas, B. C., Sharon, I., Castelle, C. J., Singh, A., Wilkins, M. J., Wrighton, K. C., Williams, K. H. and Banfield, J. F. (2015). Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature*. 523(7559): 208.

7. Buchfink, B., Xie, C. and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*. 12: 59-60.

8. Cunliffe, D. A. (1991). Bacterial nitrification in chloraminated water supplies. *Applied and Environmental Microbiology*. 57(11): 3399-3402.

9. Daims, H., Lebedeva, E. V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A. and Kirkegaard, R. H. (2015). Complete nitrification by Nitrospira bacteria. *Nature*. 528(7583): 504.

10. Daims, H., Lücker, S. and Wagner, M. (2016). A new perspective on microbes formerly known as nitrite-oxidizing bacteria. *Trends in Microbiology*. 24(9): 699-712.

11. Daims, H. and Wagner, M. (2018). Nitrospira. *Trends in Microbiology*. 26(5): 462-463.

12. De Vet, W. W. J. M., Kleerebezem, R., Van der Wielen, P. W. J. J., Rietveld, L. C. and van Loosdrecht, M. C. M. (2011). Assessment of nitrification in groundwater filters for drinking water production by qPCR and activity measurement. *Water Research*. 45(13): 4008-4018.

13. Duff, A. M., Zhang, L-M and Smith, C. J. (2017). Small-scale variation of ammonia oxidisers within intertidal sediments dominated by ammonia-oxidising bacteria

Nitrosomonas sp. *amoA* genes and transcripts. *Scientific Reports*. 7:13200. Doi: 10.1038/s41598-017-13583.

14. Emerson, D. and De Vet, W. (2015). The role of FeOB in engineered water ecosystems: a review. *Journal‑American Water Works Association*. 107(1): E47-E57.

15. Eren, A. M., Esen, Ö. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L. and Delmont, T. O. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ*. 3: e1319.

16. Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*. 17(6), pp.368-376.

17. Fowler, S. J., Palomo, A., Dechesne, A., Mines, P. D. and Smets, B. F. (2018). Comammox Nitrospira are abundant ammonia oxidizers in diverse groundwater-fed rapid sand filter communities. *Environmental Microbiology*. 20(3): 1002-1015.

18. Francis, C. A., Roberts, K. J., Beman, J. M., Santoro, A. E. and Oakley, B. B. (2005). Ubiquity and diversity of ammonia-oxidising archaea in water columns and sediments of the ocean. *Proceedings of the National Academy of Sciences*. 102(41): 14683-14688.

19. Gülay, A., Musovic, S., Albrechtsen, H. J., Al-Soud, W. A., Sorensen, 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.* doi:10.1038/ismej.2016.16

20. Gomez-Alvarez, V., Revetta, R. P and Santo Domingo, J. W. (2012). Metagenomic analysis of drinking water receiving different disinfection treatments. *Applied and Environmental Microbiology*. 78(17): 6095-6102.

21. Hall, T., Biosciences, I. and Carlsbad, C. (2011). BioEdit: an important software for molecular biology. *GERF Bulleton of Bioscience*. 2(1): 60-61.

22. Hammes, F., Berney, M., Wang, Y., Vital, M., Köster, O. and Egli, T. (2008). Flow-cytometric total bacterial cell counts as a descriptive microbiological parameter for drinking water treatment processes. *Water Research*. 42(1-2): 269-277.

23. Hoefel, D., Monis, P. T., Grooby, W. L., Andrews, S and Saint, C. P. (2005). Culture-independent techniques for rapid detection of bacteria associated with loss of chloramine residual in a drinking water system. *Applied and Environmental Microbiology*. 71(11): 6479-6488.

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

25. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M. (2015). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research*. 44(D1): D457-D462.

26. Kanehisa, M., Sato, Y. and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology*. 428: 726-731.

27. Katoh, K., Misawa, K., Kuma, K. I. and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*. 30(14): 3059-3066.

28. 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.

29. Kitzinger, K., Koch, H., Lücker, S., Sedlacek, C. J., Herbold, C., Schwarz, J., Daebeler, A., Mueller, A. J., Lukumbuzya, M., Romano, S. and Leisch, N. (2018). Characterization of the first "*Candidatus* Nitrotoga" isolate reveals metabolic versatility and separate evolution of widespread nitrite-oxidizing bacteria. *mBio*. 9(4): e01186-18.

30. Koch, H., Lucker, S., Albertsen, M., Kitzinger, K., Herbold, C., Spieck, E., Halkjaer Nielsen, P., Wagner, M. and Diams, H. (2015). Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus *Nitrospira*. *Proceedings of the National Academy of Sciences*. 112(36): 11371-11376.

31. Kraft, B., Tegetmeyer, H. E., Sharma, R., Klotz, M. G., Ferdelman, T. G., Hettich, R. L., Geelhoed, J. S. and Strous, M. (2014). The environmental controls that govern the end product of bacterial nitrate respiration. *Science*. 345(6197): 676-679.

32. Kumar, S., Stecher, G. and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular biology and evolution*. 33(7): 1870-1874.

33. Kuypers, M. M. M., Marchant, H. K and Kartal, B. (2018). The microbial nitrogen-cycling network. *Nature Reviews, Microbiology*. 16(5): 263

34. Lee, M. D. (2019). GToTree: a user-friendly workflow for phylogenomics. *BioRxiv*. p.512491.

35. Lücker, S., Wagner, M., Maixner, F., Pelletier, E., Koch, H., Vacherie, B., Rattei, T., Damsté, J.S.S., Spieck, E., Le Paslier, D. and Daims, H. (2010). A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-

oxidizing bacteria. Proceedings of the National Academy of Sciences. 107(30): 13479-13484.

36. Maixner, F., Noguera, D. R., Anneser, B., Stoecker, K., Wegl, G., Wagner, M. and Daims, H. (2006). Nitrite concentration influences the population structure of Nitrospira-like bacteria. *Environmental Microbiology*. 8(8): 1487-1495.

37. Matsen, F. A., Kodner, R. B. and Armbrust, E. V. (2010). pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics*. 11(1): 538.

38. McCoy, S. T. and VanBriesen, J. M. (2014). Comparing spatial and temporal diversity of bacteria in a chlorinated drinking water distribution system. *Environmental Engineering Science*. 31(1): 32-41.

39. Moradi, S., Liu, S., Chow, C.W., van Leeuwen, J., Cook, D., Drikas, M. and Amal, R. (2017). Developing a chloramine decay index to understand nitrification: A case study of two chloraminated drinking water distribution systems. *Journal of Environmental Sciences*. 57: 170-179.

40. Moreno-Vivián, C., Cabello. P., Martínez-Luque, M., Blasco, R. and Castillo, F. (1999). Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. *Journal of Bacteriology*. 181(21): 6573-84.

41. Nawrocki, E. P., Kolbe, D. L. and Eddy, S. R. (2009). Infernal 1.0: inference of RNA alignments. *Bioinformatics*. 25(10): 1335-1337.

42. Nicol, G. W. and Schleper, C. (2006). Ammonia-oxidising Crenachaeota: important players in the nitrogen cycle? *Trends in Microbiology*. 12(5): 207-212.

43. Norton, C. D. and LeChevallier, M. W. (1997). Chloramination: its effect on distribution system water quality. *Journal-American Water Works Association*. 89(7): 66-77.

44. Nurk, S., Meleshko, D., Korobeynikov, A. and Pevzner, P.A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Research*. 27(5): 824-834.

45. Palatinszky, M., Herbold, C., Jehmlich, N., Pogoda, M., Han, P., von Bergen, M., Lagkouvardos, I., Karst, S.M., Galushko, A., Koch, H and Berry, D. (2015). Cyanate as an energy source for nitrifiers. *Nature*. 524(7563): 105.

46. Palomo, A., Pedersen, A. G., Fowler, S. J., Dechesne, A., Sicheritz-Pontén, T. and Smets, B. F. (2018). Comparative genomics sheds light on niche differentiation and the evolutionary history of comammox Nitrospira. *The ISME journal*. p.1.

47. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*. doi:10.1101/gr.186072.114.

48. Pester, M., Schleper, C. and Wagner, M. (2011). The *Thaumarchaeota*: an emerging view of their phylogeny and ecophysiology. *Current Opinion in Microbiology*. 14(3): 300-306.

49. Philippot, L. (2005). Tracking nitrate reducers and denitrifiers in the environment. *Biochemical Society Transactions*. (33): 200-204.

50. Pintar, K. D. M. and Slawson, R. M. (2003). Effect of temperature and disinfection strategies on ammonia-oxidising bacteria in a bench-scale drinking water distribution system. *Water Research*. 37: 1805-1817.

51. 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.

52. 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 ASM*. 5(3): e01135-14.

53. Pinto, A., Marcus, D. N., Ijaz, U. Z., Bautista-de los Santos, Q. M., Dick, G. J. and Raskin, L. (2015). Metagenomic evidence for the presence of comammox *Nitrospira*-like bacteria in drinking water system. *mSphere*. 1(1): e00054-15.

54. Potgieter, S., Pinto, A., Sigudu, M., Ncube. E. J., du Preez, H. and Venter, S. N. (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.

55. Prest, E. I., Weissbrodt, D. G., Hammes, F., van Loosdrecht, M. C. M and Vrouwenvelder, J. S. (2016). Long-term bacterial dynamics in a full scale drinking water distribution system. *PLoS ONE*. DOI: 10.1371/journal.pone.0164445.

56. Regan, J. M., Harrington, G. W., Baribeau, H., De Leon, R. and Noguera, D. R. (2003). Diversity of nitrifying bacteria in full-scale chloraminated distribution systems. *Water Research*. 37(1): 197-205.

57. Richardson, D. J., Berks, B. C., Russell, D. A., Spiro, S. and Taylor, C. J. (2001). Functional, biochemical and genetic diversity of prokaryotic nitrate reductases. *Cellular and Molecular Life Sciences CMLS*. 58(2): 165-178.

58. Rivett, M. O., Buss, S. R., Morgan, P., Smith, J. W. and Bemment, C. D. (2008). Nitrate attenuation in groundwater: a review of biogeochemical controlling processes. *Water Research*. 42(16): 4215-4232.

59. Rodriguez-R, L. M. and Konstantinidis, K. T. (2013). Nonpareil: a redundancy-based approach to assess the level of coverage in metagenomic datasets. *Bioinformatics*. 30(5): 629-635.

60. Rodriguez-R, L. M. and Konstantinidis, K. T. (2014). Bypassing cultivation to identify bacterial species. *Microbe*. 9(3): 111-118.

61. 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*.

62. Sawade, E., Monis, P., Cook, D. and Drikas, M. (2016). Is nitrification the only cause of microbiologically induced chloramine decay?. *Water Research*. 88: 904-911.

63. Schleper, C. (2010). Ammonia oxidation: different niches for bacteria and archaea? *The ISME Journal*. 4(9): 1092.

64. 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. and 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): 7537-7541.

65. Shaw, J. L. A., Monis, P., Weyrich, L. S., Sawade, E., Drikas, M. and Cooper, A. J. (2015). Using amplicon sequencing to characterize and monitor bacterial diversity in drinking water distribution systems. *Applied and Environmental Microbiology*. 81(18): 6463-6473.

66. Smith, C. J., Nedwell, D. B., Dong, L. F. and Osborn, A. M. (2007). Diversity and abundance of nitrate reductase genes (*narG* and *napA*), nitrite reductase genes (*nirS* and *nrfA*), and their transcripts in estuarine sediments. *Applied and Environmental Microbiology*. 73(11): 3612-3622.

67. Smith, C. J., Dong, L. F., Wilson, J., Scott, A., Osborn, A. M. and Nedwell, D. B. (2016). Seasonal variation in denitrification and dissimilatory nitrate reduction to ammonia process rates and corresponding key functional genes along an estuarine nitrate gradient. *Frontiers in Microbiology*. doi.org/10.3389/fmicb.2015.00542.

68. Stein, L. Y. and Klotz, M. G. (2011). Nitrifying and denitrifying pathways of methanotrophic bacteria. p 1826-1831.

69. Urakawa, H., Martens-Habbena, W. and Stahl, D. A. (2010). High abundance of ammonia-oxidising Archaea in coastal waters, determined by using a modified DNA extraction method. *Applied and Environmental Microbiology*. 76(7): 2129-2135.

70. Ushiki, N., Jinno, M., Fujitani, H., Suenaga, T., Terada, A. and Tsuneda, S. (2017). Nitrite oxidation kinetics of two *Nitrospira* strains: The quest for competition and ecological niche differentiation. *Journal of Bioscience and Bioengineering*. 123(5): 581-589.

71. van der Kooij, D. (2002). Assimilable organic carbon (AOC) in treated water: determination and significance. In: Bitton, G. (Ed.), Encyclopedia of Environmental Microbiology. Wiley, Hoboken, NJ, USA. pp: 312–327.

72. Van der Wielen, P. W. J. J., Voost, S and van der Kooij, D. (2009). Ammonia-oxidising bacteria and archaea in groundwater treatment and drinking water distribution systems. *Applied and Environmental Microbiology*. 75(14): 4687-4695.

73. van Kessel, M. A., Speth, D. R., Albertsen, M., Nielsen, P. H., den Camp, H. J. O., Kartal, B., Jetten, M. S. and Lücker, S. (2015). Complete nitrification by a single microorganism. *Nature*. 528(7583): 555.

74. Vikesland, P. J., Ozekin, K and Valentine, R. L. (2001). Monochloramine decay in model and distribution system waters. *Water Research*. 35(7): 1766-1776.

75. Wang, H., Proctor, C. R., Edwards, M. A., Pryor, M., Santo Domingo, J. W., Ryu, H., Camper, A. K., Olson, A and Pruden, A. (2014). Microbial community response to chlorine conversion in a chloraminated drinking water distribution system. *Environmental Science and Technology*. 48(18): 10624-10633.

76. Wang, Y., Ma, L., Mao, Y., Jiang, X., Xia, Y., Yu, K., Li, B. and Zhang, T. (2017). Comammox in drinking water systems. *Water Research*. 116: 332-341.

77. Wolfe, R. L., Lieu, N. I., Izaguirre, G. and Means, E. G. (1990). Ammonia-oxidising bacteria in a chloraminated distribution system: seasonal occurrence, distribution and disinfectant resistance. *Applied and Environmental Microbiology*. 56(2): 451-462.

78. Zhang, Y., Love, N. and Edwards, M. (2009). Nitrification in drinking water systems. *Critical Reviews in Environmental Science and Technology*. 39: 153-208.

79. Zlatanovic, L., van der Hoek, J. P. and Vreeburg, J. H. G. (2017). An experimental study on the influence of water stagnation and temperature change on water quality in a full-scale domestic drinking water system. *Water Research*. 123: 761-772.

Figures
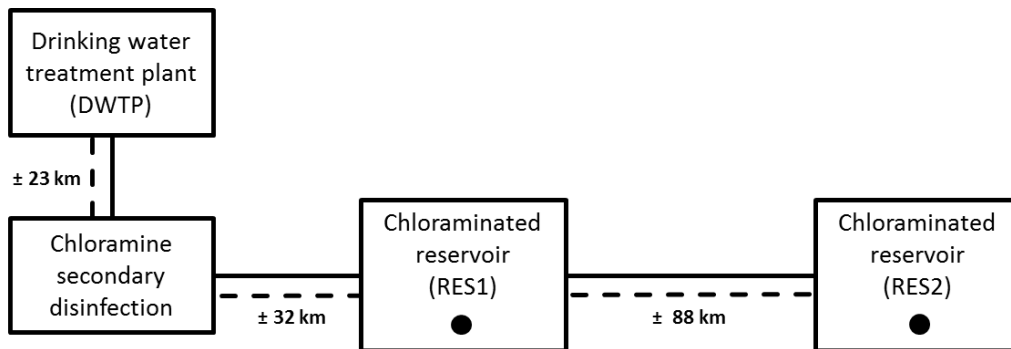
Figure 1: Simplified schematic showing the layout of the DWDS and the sample locations [reservoir 1 (RES1) and reservoir 2 (RES2) indicated in the figure as black circles]. Approximate distances between locations indicated in the figure as the dotted line.
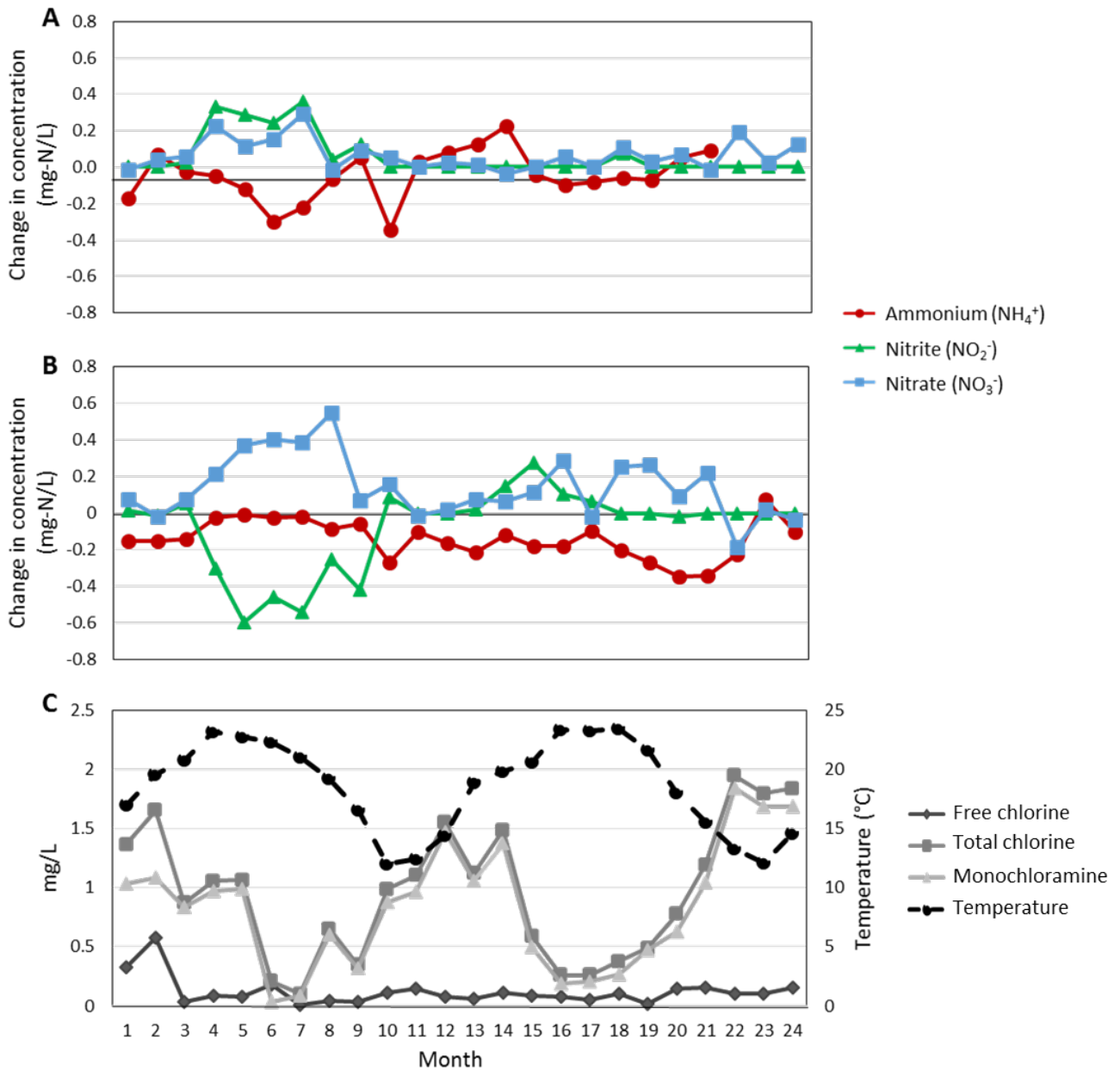
Figure 2: Change in nitrogen species concentration [i.e., ammonium (red circles), nitrite (green triangles) and nitrate (blue squares)] (A) before and after reservoir 1 (RES1) and (B) before and after reservoir 2 (RES2) over the two year sampling period. (C) Average concentrations of disinfectant residuals [i.e., free chlorine (diamonds), total chlorine (squares) and monochloramine (triangles)] together with average temperature (dashed black line with circles) across RES1 and RES2.
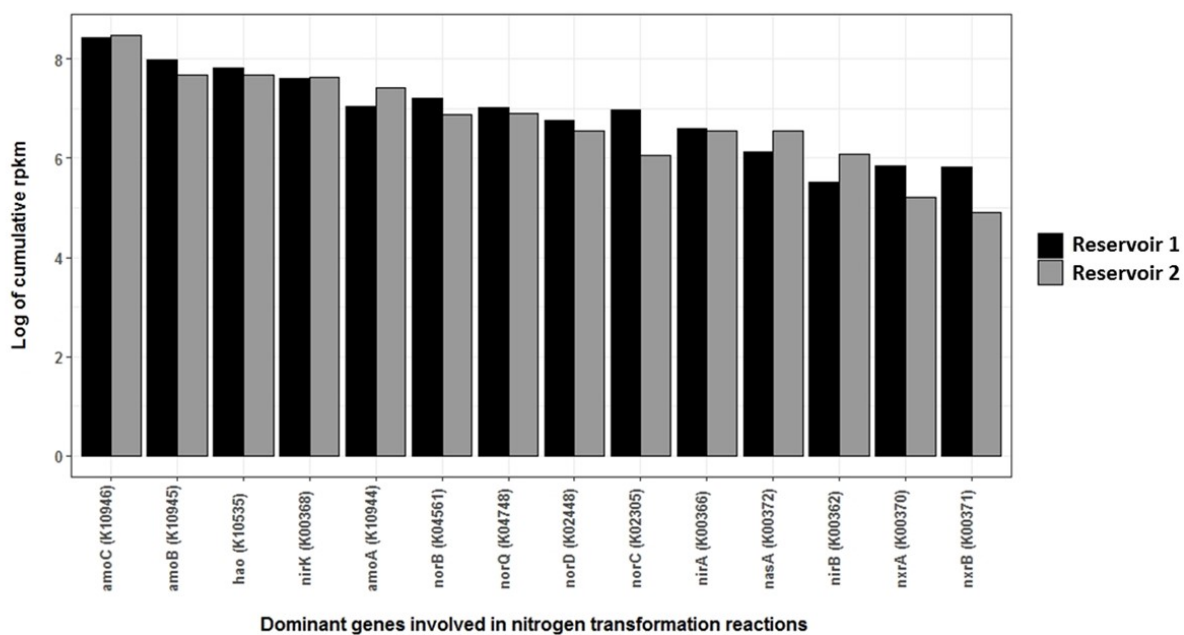
Figure 3: Log transformed cumulative coverage (rpkm) of the dominant genes identified to be involved in the nitrogen cycle (i.e. genes with a cumulative coverage of >100 rpkm in both reservoirs).

Figure 4: Cumulative coverage (rpkm) of the dominant genes identified to be involved in the nitrogen cycle (i.e. genes with a cumulative coverage of >100 rpkm in both reservoirs) across all reservoir samples. This figure is complemented by Figure C5, showing the number of genes identified for each function.

A

- Nitrosomonas sp. AL212 (WP_013646792)
- **NODE_139662**
- Nitrosomonas ureae (SEF78994)
- Nitrosomonas marina (APZ84151)
- Nitrosomonas sp. Is79A3 (WP_013964654)
- **NODE_6918**
- **NODE_289595**
- Nitrosomonas oligotropha (APZ84147)
- **NODE_200396**
- **NODE_21034**
- **NODE_43416**
- Nitrosomonas cryotolerans (AAG60667)
- Nitrosospira multiformis (AAC25057)
- Nitrosospira briensis (AAB38709)
- Nitrosovibrio tenuis (AAB38710)
- **NODE_9011**
- **NODE_49621**
- Nitrosomonas communis (APZ84141)
- Nitrosomonas nitrosa (SFM93401)
- Nitrosomonas halophila (SDZ04661)
- Nitrosomonas mobilis (SCZ85183)
- Nitrosomonas eutropha (SEJ33789)
- Nitrosomonas mobilis (AAC31361)
- Nitrosomonas europaea (KXK49931)
- Ca. Nitrospira nitrosa (WP_090742150)
- Ca. Nitrospira inopinata (CUQ66826)
- Ca. Nitrospira nitrificans (CUS38627)
- Nitrospira sp. Ga0074138
- Nitrospira SG bin1 (OQW38018)
- Nitrospira sp. CMX UWM
- Nitrospira comammox A2
- Nitrospira SG bin2 (OQW37964)
- Nitrospira ST bin4 (OQW57195)
- **NODE 6011**
- Methylosinus sp. LW4 (WP 018265987)
- Methylocystis parvus (WP 026016563)
- Nitrosococcus watsonii (AAF03938)
- Nitrosococcus oceani (AAC25091)
- Methylococcus capsulatus (AAB49821)
- Methylomicrobium buryatense (WP 017841993)
- Methylomonas methanica (WP 013817026)
- Methylobacter luteus (WP 027159170)

0.050

B

- Ca. Nitrospira sp. ST-bin5 (OQW66184)
- **NODE_6313**
- **NODE_263397**
- **NODE_50179**
- **NODE_404490**
- **NODE_16282**
- **NODE_2534 (C56)**
- **NODE_3422 (C56)**
- **NODE_265263**
- **NODE_269148**
- **NODE_68 (C51)**
- **NODE_19179**
- Nitrospira lenta (SPP64235)
- Nitrospira moscoviensis (WP_053381689)
- **NODE_328673**
- Ca. Nitrospira nitrosa (CUS33249)
- **NODE_44928**
- Ca. Nitrospira inopinata (WP_062483509)
- **NODE_18132**
- **NODE_174022**
- Ca. Nitrospira nitrificans (CUS36764)
- Nitrospira nitrificans
- Nitrospira defluvii (WP_013249749)
- Nitrospira japonica (WP 080885591)
- Nitrospina sp. (KMP11423)
- Nitrospina gracilis (WP_042250442)
- Ca. Scalindua brodae (KHE93157)
- **NODE_199644**
- Ca. Kuenenia stuttgartiensis (CAJ72445)
- Ca. Brocadia sapporoensis (WP_070066136)
- Ca. Brocadia fulgida (KKO18748)
- Ca. Jettenia caeni (WP_007220240)
- Ca. Brocadia sinica (KXK29028)
- Ca. Brocadia caroliniensis (OOP55454)
- Ca. Methanoperedens Nitroreducens (KCZ70344)
- Ca. Nitrotoga fabula (SPS06751)
- Ca. Methanoperedens sp. BLZ1 (KPQ44267)
- Nitrolancea hollandica (AFN37205)
- Nitrococcus mobilis (ZP_01125872)
- Nitrobacter hamburgensis (YP_578186)
- Nitrobacter winogradskyi (YP_318678)
- Nitrobacter alkalicus (AAQ14967)
- Nitrobacter vulgaris (AAQ14969)
- Nitrobacter sp. Nb-311A (ZP_01048070)

0.05

211

Figure 5: Phylogenetic placement of (A) ammonia monooxygenase, subunit A (*amoA*) and (B) nitrite oxidoreductase, subunit A (*nxrA*) phylogenies. Both Maximum Likelihood phylogenetic trees were constructed based on amino acid sequences from contigs identified as the respective genes in the metagenomic dataset. Contigs identified from this study are in bold. Reference trees were constructed with bootstrap analysis of 1000 replicates. Bootstrap values are indicated as percentages and values below 50 were excluded.

A

B

	amoA (K10944)	amoB (K10945)	amoC (K10946)	hao (K10535)	nxrA (K00370)	nxrB (K00371)	nirA (K00366)	nirB (K00362)	nasA (K00372)	nirK (K00368)	norB (K04561)	norC (K02305)	norD (K02448)	norQ (K04748)

213

Figure 6: (A) Phylogenomic tree showing the genome-level evolutionary inference of the 47 MAGs constructed from the metagenomic dataset and (B) the corresponding presence of dominant genes associated with the nitrogen cycle identified in each MAG. (Green indicates those genes involved in nitrification, blue indicates genes involved in nitrate and nitrite reduction and red indicates genes involved in nitric oxide formation and reduction).
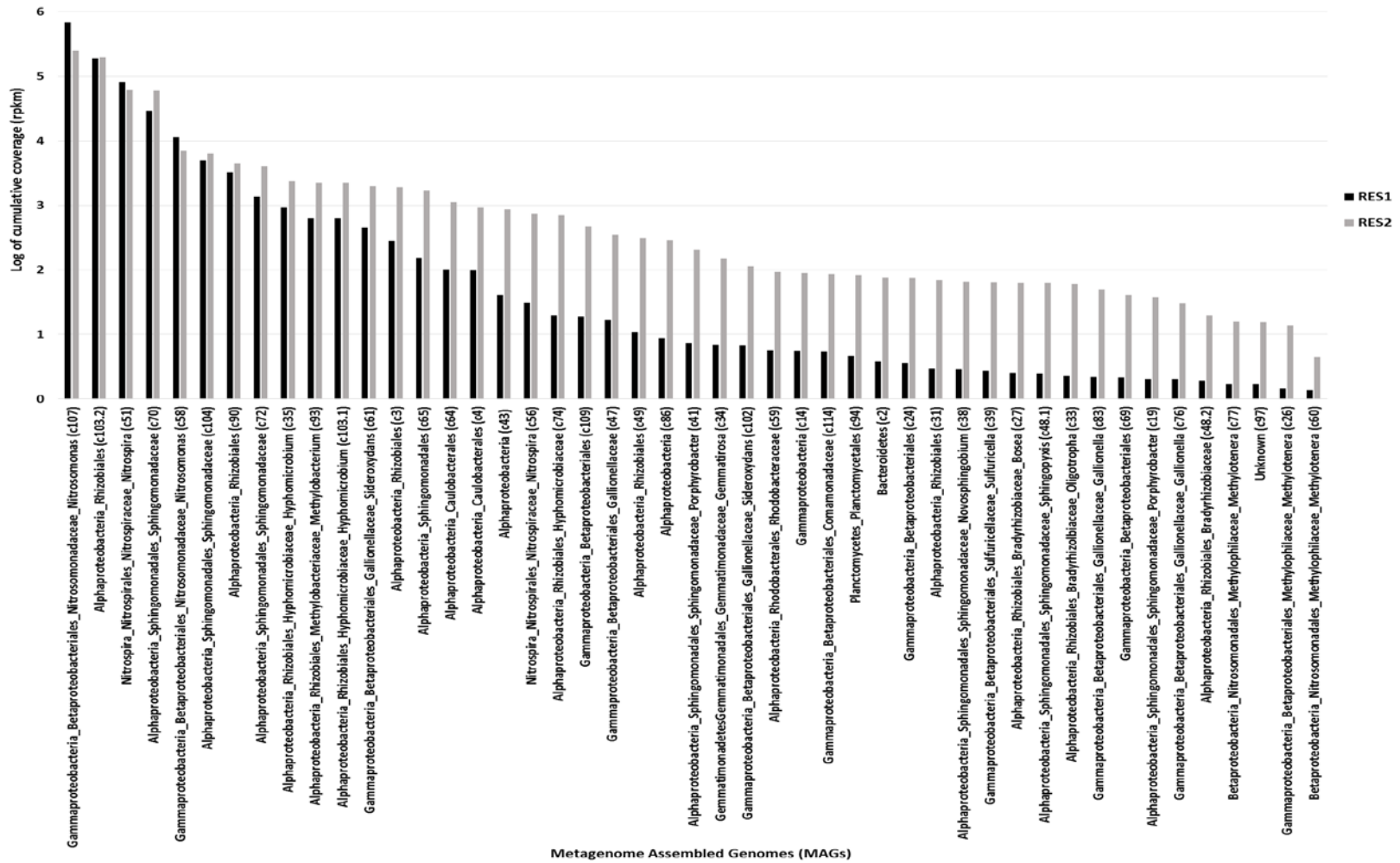
Figure 7: Log transformed cumulative coverage (rpkm) of the 47 reconstructed Metagenome Assembled Genomes (MAGs) for both reservoirs.

Figure 8: A schematic overview of the dominant genes and MAGs involved in the major reactions of the nitrogen cycle in both chloraminated reservoirs. Reactions involved in nitrification are indicated in green, reactions involved in assimilatory nitrate reduction are indicated in blue and the reactions involving the formation and reduction of nitric oxide are indicated in red.

Appendix C: Chapter 4 supplementary information

Supplementary tables

Table C1A: Chemical monitoring data for the duration of the study obtained from the utility for reservoir 1 (RES1)

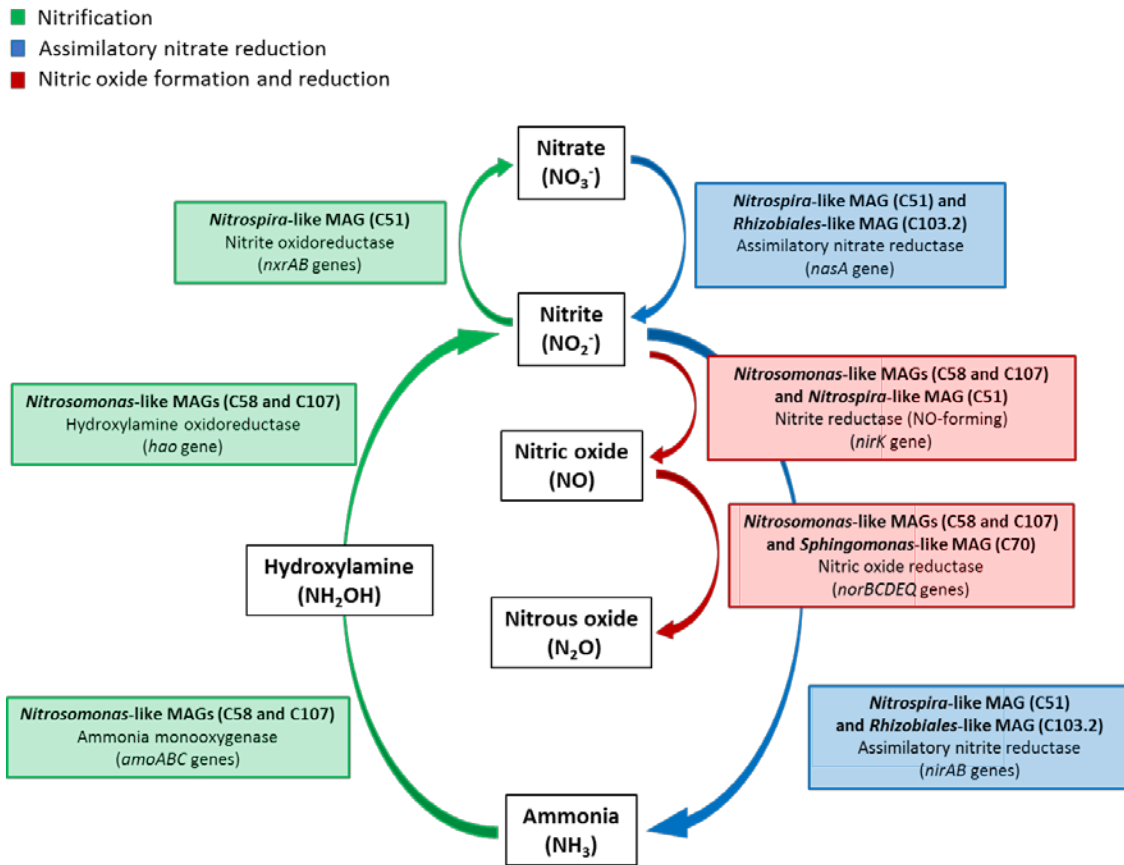|	Month	Free and available chlorine (mg/l)	Total chorine (mg/l)	Monochloramine (mg/l)	Total residual Chorine (mg/l)	Temp (°C)	Δ ammonium (mg/l)	Δ nitrite (mg/l)	Δ nitrate (mg/l)
1	7-Oct-2014	0.53	2.09	1.56	2.12	16.1	-0.17	0.00	-0.02
2	4-Nov-2014	0.3	2.2	1.9	2.2	18.5	0.07	0.00	0.04
3	2-Dec-2014	0.03	0.74	0.71	0.81	22.3	-0.02	0.02	0.06
4	13-Jan-2015	0.07	1.94	1.87	1.94	23.4	-0.05	0.33	0.23
5	3-Feb-2015	0.11	2.07	1.96	2.09	22.7	-0.12	0.29	0.12
6	3-Mar-2015	0.34	0.36	0.02	0.36	22.2	-0.30	0.24	0.16
7	14-Apr-2015	0	0.15	0.15	0.18	20.8	-0.22	0.36	0.29
8	5-May-2015	0.06	1.26	1.2	1.28	18.4	-0.06	0.04	-0.01
9	2-Jun-2015	0.06	0.67	0.61	0.68	16.8	0.05	0.13	0.09
10	7-Jul-2015	0.18	1.35	1.17	1.68	10.5	-0.34	0.00	0.05
11	4-Aug-2015	0.09	0.96	0.87	0.95	12	0.03	0.00	0.00
12	1-Sep-2015	0.11	1.66	1.55	1.61	14.7	0.08	0.00	0.03
13	6-Oct-2015	0.01	1.42	1.41	1.42	19	0.12	0.00	0.02
14	3-Nov-2015	0.12	1.69	1.57	1.74	19.5	0.23	0.00	-0.04
15	1-Dec-2015	0.06	0.11	0.05	0.13	20.7	-0.04	0.00	0.01

16	5-Jan-2016	0.03	0.1	0.07	0.14	22.7	-0.10	0.00	0.06
17	2-Feb-2016	0	0.33	0.33	0.35	23.2	-0.08	0.00	0.00
18	8-Mar-2016	0.06	0.27	0.21	0.3	23.5	-0.06	0.08	0.11
19	5-Apr-2016	0.03	0.48	0.45	0.59	21.7	-0.07	0.00	0.03
20	10-May-2016	0.13	1.07	0.94	1.07	18.1	0.06	0.00	0.07
21	7-Jun-2016	0.24	0.84	0.6	0.97	15.2	0.09	0.00	-0.02
22	5-Jul-2016	0.15	2.11	1.96	2.11	12.8	-	0.00	0.20
23	2-Aug-2016	0.09	1.84	1.75	1.99	11.9	-	0.00	0.02
24	30-Aug-2016	0.21	1.94	1.73	1.86	14.2	-	0.00	0.13

Table C1B: Chemical monitoring data for the duration of the study obtained from the utility for reservoir 2 (RES2)

	Month	Free and available chlorine (mg/l)	Total chorine (mg/l)	Monochloramine (mg/l)	Total residual Chorine (mg/l)	Temp (°C)	Δ ammonium (mg/l)	Δ nitrite (mg/l)	Δ nitrate (mg/l)
1	10-Oct-2014	0.13	0.64	0.51	0.98	17.9	-0.15	0.02	0.08
2	7-Nov-2014	0.85	1.11	0.26	1.13	20.6	-0.15	-0.02	-0.02
3	1-Dec-2014	0.04	1	0.96	1	19.3	-0.14	0.05	0.07
4	16-Jan-2015	0.1	0.17	0.07	0.18	23	-0.02	-0.30	0.21
5	6-Feb-2015	0.04	0.06	0.02	0.09	22.8	-0.01	-0.60	0.37
6	6-Mar-2015	0.02	0.06	0.04	0.08	22.4	-0.03	-0.46	0.40
7	17-Apr-2015	0.02	0.05	0.03	0.07	21.2	-0.02	-0.54	0.39
8	8-May-2015	0.03	0.04	0.01	0.05	20	-0.09	-0.25	0.55
9	5-Jun-2015	0	0.02	0.02	0.03	16.4	-0.06	-0.42	0.07
10	10-Jul-2015	0.04	0.62	0.58	0.65	13.5	-0.27	0.09	0.16
11	7-Aug-2015	0.2	1.25	1.05	1.23	12.9	-0.10	0.00	-0.01
12	4-Sep-2015	0.05	1.44	1.39	1.47	14.2	-0.16	0.00	0.02
13	9-Oct-2015	0.11	0.82	0.71	1.21	18.7	-0.21	0.02	0.08
14	6-Nov-2015	0.11	1.28	1.17	1.35	20.1	-0.12	0.15	0.07
15	4-Dec-2015	0.12	1.07	0.95	1.13	20.5	-0.18	0.28	0.12
16	8-Jan-2016	0.12	0.42	0.3	0.42	24	-0.18	0.10	0.29
17	5-Feb-2016	0.11	0.19	0.08	0.23	23.4	-0.10	0.06	-0.02
18	11-Mar-2016	0.15	0.48	0.33	0.51	23.4	-0.21	0.00	0.26

19	8-Apr-2016	0	0.5	0.5	0.51	21.6	-0.27	0.00	0.27
20	13-May-2016	0.16	0.48	0.32	0.49	18	-0.35	-0.02	0.10
21	10-Jun-2016	0.07	1.55	1.48	1.57	15.9	-0.34	0.00	0.22
22	8-Jul-2016	0.05	1.78	1.73	1.75	13.7	-0.23	0.00	-0.18
23	5-Aug-2016	0.12	1.75	1.63	1.86	12.3	0.07	0.00	0.02
24	2-Sep-2016	0.1	1.74	1.64	1.69	14.9	-0.10	0.00	-0.03

Table C2: Outline and description of the chloraminated reservoir samples collected over the 2-year study period together with the resulting DNA concentrations

	Sample Name	Month	Year	Qubit DNA concentration (ng/ul)
1	RES1_4	January	2015	35.0
2	RES1_5	February	2015	23.2
3	RES1_6	March	2015	13.3
4	RES1_17	February	2016	40.2
5	RES1_18	March	2016	19.7
6	RES1_19	April	2016	12.6
7	RES1_20	May	2016	10.6
8	RES1_21	June	2016	6.92
9	RES2_1	October	2014	13.8
10	RES2_5	February	2015	14.6
11	RES2_6	March	2015	42.8
12	RES2_7	April	2015	7.84
13	RES2_8	May	2015	23.3
14	RES2_10	July	2015	17.6
15	RES2_17	February	2016	11.3
16	RES2_18	March	2016	20.8
17	RES2_19	April	2016	30.8
18	RES2_20	May	2016	24.6

Table C3: The final number of reads mapped to assembly

	Sample	Number of reads	Mapped reads to scaffolds > 500 bp	% Mapped reads
1	RES1_4	14 249 511	14 195 821	99.62
2	RES1_5	13 019 437	12 929 842	99.31
3	RES1_6	12 747 382	12 673 029	99.42
4	RES1_17	14 647 502	14 573 146	99.49
5	RES1_18	18 488 103	18 376 301	99.40
6	RES1_19	11 184 973	11 116 758	99.39
7	RES1_20	6 855 217	6 753 636	98.52
8	RES1_21	16 944 215	16 809 348	99.20
9	RES2_1	12 231 070	11 382 585	93.06
10	RES2_5	11 677 750	11 487 402	98.37
11	RES2_6	6 919 446	6 758 814	97.68
12	RES2_7	11 590 760	11 452 070	98.80
13	RES2_8	13 020 429	12 809 835	98.38
14	RES2_10	13 802 699	13 303 643	96.38
15	RES2_17	13 688 540	13 501 850	98.64
16	RES2_18	13 259 038	13 173 737	99.36
17	RES2_19	21 940 278	21 877 986	99.72
18	RES2_20	12 542 819	12 483 217	99.52

Table C4: Summary of the genes identified (including KEGG orthology numbers) involved in nitrogen transformation pathways in chloraminated reservoir samples and their cumulative coverage (rpkm) in each reservoir

K number	Enzyme	Enzyme name	Reaction	Process	RES1*	RES2*
K10944	*amoA*	Ammonia monooxygenase, subunit A	$NH_4^+ + O_2 + 2e^- \rightarrow NH_2OH + H_2O$	Ammonia oxidation - Nitrification	1143.07	1671.50
K10945	*amoB*	Ammonia monooxygenase, subunit B	$NH_4^+ + O_2 + 2e^- \rightarrow NH_2OH + H_2O$	Ammonia oxidation - Nitrification	2922.21	2165.83
K10946	*amoC*	Ammonia monooxygenase, subunit C	$NH_4^+ + O_2 + 2e^- \rightarrow NH_2OH + H_2O$	Ammonia oxidation - Nitrification	4625.20	4807.07
K10535	*hao*	Hydroxylamine oxidoreductase	$NH_2OH \rightarrow NO + 3e^- + 3H^+$	Hydroxylamine oxidation - Nitrification	2486.24	2169.57
K05601	*hcp*	Hydroxylamine reductase	$NH_2OH + H_2O \rightarrow NH_4^+ + O_2 + 2e^-$	Hydroxylamine reduction - Nitrification	6.76	53.03
K00370	*nxrA*	Nitrite oxidoreductase, alpha subunit	$NO_2^- + H_2O \rightarrow NO_3^- + 2e^- + 2H^+$	Nitrite oxidation - Nitrification	346.72	183.14
K00371	*nxrB*	Nitrite oxidoreductase, beta subunit	$NO_2^- + H_2O \rightarrow NO_3^- + 2e^- + 2H^+$	Nitrite oxidation - Nitrification	335.18	135.21
K00370	*narG*	Respiratory nitrate reductase, alpha subunit	$NO_3^- + 2e^- + 2H^+ \rightarrow NO_2^- + H_2O$	Dissimilatory nitrate reduction	66.10	278.95
K00371	*narH*	Respiratory nitrate reductase, beta subunit	$NO_3^- + 2e^- + 2H^+ \rightarrow NO_2^- + H_2O$	Dissimilatory nitrate reduction	35.15	205.32
K00373	*narJ*	Respiratory nitrate reductase, delta subunit (molybdenum cofactor assembly chaperone)	$NO_3^- + 2e^- + 2H^+ \rightarrow NO_2^- + H_2O$	Dissimilatory nitrate reduction	18.09	159.90
K00374	*narI*	Respiratory nitrate reductase, gamma subunit	$NO_3^- + 2e^- + 2H^+ \rightarrow NO_2^- + H_2O$	Dissimilatory nitrate reduction	23.94	182.24
K00372	*nasA*	Assimilatory nitrate reductase, large catalytic subunit	$NO_3^- + 2e^- + 2H^+ \rightarrow NO_2^- + H_2O$	Assimilatory nitrate reduction	455.05	701.03
K00362	*nirB*	Nitrite reductase (NAD(P)H), large subunit	$NO_2^- + 6e^- + 8H^+ \rightarrow NH_4^+ + 2H_2O$	Dissimilatory nitrite reduction	248.74	437.86
K00363	*nirD*	Nitrite reductase (NAD(P)H), small	$NO_2^- + 6e^- + 8H^+ \rightarrow NH_4^+ + 2H_2O$	Dissimilatory nitrite reduction	94.70	310.78

			subunit			
K03385	*nrfA*	Cytochrome c-552	$NO_2^- + 6e^- + 8H^+ \rightarrow NH_4^+ + 2H_2O$	Dissimilatory nitrite reduction	3.98	32.21
K00366	*nirA*	Ferredoxin-nitrite reductase	$NO_2^- + 6e^- + 8H^+ \rightarrow NH_4^+ + 2H_2O$	Assimilatory nitrite reduction	729.37	692.10
K00368	*nirK*	Nitrite reductase (NO-forming)	$NO_2^- + e^- + 2H^+ \rightarrow NO + H_2O$	Nitrite reduction - Nitric oxide (NO) forming	2016.40	2043.11
K02448	*norD*	Nitric oxide reductase, D protein	$2NO + 2e^- + 2H^+ \rightarrow N_2O + H_2O$	Nitric oxide reduction	856.37	699.77
K04561	*norB*	Nitric oxide reductase, subunit B	$2NO + 2e^- + 2H^+ \rightarrow N_2O + H_2O$	Nitric oxide reduction	1363.59	971.78
K02305	*norC*	Nitric oxide reductase, subunit C	$2NO + 2e^- + 2H^+ \rightarrow N_2O + H_2O$	Nitric oxide reduction	1060.40	427.07
K04748	*norQ*	Nitric oxide reductase, Q protein	$2NO + 2e^- + 2H^+ \rightarrow N_2O + H_2O$	Nitrous-oxide reduction	1121.85	996.92
K00376	*nosZ*	Nitrous-oxide reductase	$N_2O + 2e^- 2H^+ \rightarrow N_2 + H_2O$	Nitrogenase - Nitrogen fixation	62.22	85.89
K02586	*nifD*	Nitrogenase molybdenum-iron protein, alpha chain	$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i$	Nitrogenase - Nitrogen fixation	32.06	8.58
K02591	*nifK*	Nitrogenase molybdenum-iron protein, beta chain	$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i$	Nitrogenase - Nitrogen fixation	30.73	4.64
K02588	*nifH*	Nitrogenase iron protein	$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i$	Nitrogenase - Nitrogen fixation	31.56	9.82

* Cumulative reads per kilobase million (rpkm) of each gene in each reservoir

Table C5: Reference genomes of nitrifying organisms

	Reference name	Accession number	Nitrifier type
1	*Ca*. Nitrosoarchaeum sp.	NSIR01000002	Ammonia oxidizing archaea
2	*Ca*. Nitrosoarchaeum koreensis	NZ AFPU01000001	Ammonia oxidizing archaea
3	*Ca*. Nitrosopumilus adriaticus	NZ CP011070	Ammonia oxidizing archaea
4	*Ca*. Nitrosopumilus koreensis	CP003842	Ammonia oxidizing archaea
5	*Ca*. Nitrosopumilus piranensis	CP010868	Ammonia oxidizing archaea
6	*Ca*. Nitrosopumilus sp.	NC 018656	Ammonia oxidizing archaea
7	*Ca*. Nitrososphaera evergladensis	CP007174	Ammonia oxidizing archaea
8	*Ca*. Nitrososphaera gargensis	CP002408	Ammonia oxidizing archaea
9	*Cenarchaeum symbiosum*	DP000238	Ammonia oxidizing archaea
10	*Nitrosopumilus maritimus*	CP000866	Ammonia oxidizing archaea
11	*Nitrososphaera viennensis*	NZ CP007536	Ammonia oxidizing archaea
12	*Nitrosococcus halophilus*	NC 013960	Ammonia oxidizing bacteria
13	*Nitrosococcus oceani*	NC 007484	Ammonia oxidizing bacteria
14	*Nitrosococcus watsonii*	NC 014315	Ammonia oxidizing bacteria
15	*Nitrosomonas aestuarii*	NZ FOSP01000116	Ammonia oxidizing bacteria
16	*Nitrosomonas communis*	NZ CP011451	Ammonia oxidizing bacteria
17	*Nitrosomonas cryotolerans*	NZ FSRO01000002	Ammonia oxidizing bacteria
18	*Nitrosomonas europaea*	NC 004757	Ammonia oxidizing bacteria
19	*Nitrosomonas eutropha*	NC 008344	Ammonia oxidizing bacteria
20	*Nitrosomonas halophila*	NZ FNOY01000156	Ammonia oxidizing bacteria
21	*Nitrosomonas marina*	NZ FOCP01000068	Ammonia oxidizing bacteria
22	*Nitrosomonas mobilis*	NZ FMWO01000108	Ammonia oxidizing bacteria
23	*Nitrosomonas nitrosa*	NZ FOUF01000076	Ammonia oxidizing bacteria
24	*Nitrosomonas oligotropha*	NZ FNOE01000080	Ammonia oxidizing bacteria
25	*Nitrosomonas ureae*	NZ CP013341	Ammonia oxidizing bacteria
26	*Nitrosospira briensis*	NZ CP012371	Ammonia oxidizing bacteria
27	*Nitrosospira lacus*	NZ CP021106	Ammonia oxidizing bacteria
28	*Nitrosospira multiformis*	NC 007614	Ammonia oxidizing bacteria
29	*Ca*. Nitrospira inopinata	NZ LN885086	Comammox bacteria
30	*Ca*. Nitrospira nitrificans	NZ CZPZ01000001	Comammox bacteria
31	*Ca*. Nitrospira nitrosa	NZ CZQA01000001	Comammox bacteria
32	*Nitrospira* sp.	LNDU01000039	Comammox bacteria
33	*Ca*. Nitrospira defluvii	NC 014355	Nitrite oxidizing bacteria
34	*Nitrobacter vulgaris*	NZ MWPQ01000040	Nitrite oxidizing bacteria
35	*Nitrobacter hamburgensis*	NC 007964	Ammonia oxidizing bacteria

36	*Nitrobacter winogradskyi*	NC 007406	Ammonia oxidizing bacteria
37	*Nitrospira japonica*	NZ LT828648	Nitrite oxidizing bacteria
38	*Nitrospira lenta*	OUNR01000001	Nitrite oxidizing bacteria
39	*Nitrospira moscoviensis*	NZ CP011801	Nitrite oxidizing bacteria
40	*Ca*. Brocadia caroliniensis	AYTS01000035	Anammox bacteria
41	*Ca*. Brocadia fulgida	LAQJ01000121	Anammox bacteria
42	*Ca*. Brocadia sapporoensis	NZ MJUW02000122	Anammox bacteria
43	*Ca*. Jettenia caeni	NZ BAFH01000003	Anammox bacteria
44	*Ca*. Kuenenia stuttgartiensis	NZ LT934425	Anammox bacteria

Table C6: Characteristics of constructed Metagenome Assembled Genomes (MAGs)

MAG	Taxonomy[a]	MiGA	Completeness %	Contamination %	Quality %	G+C content %	Number of contigs	Genetic size (Mb)	Largest contig (bp)	N50	Number of ORF's	Number of genes annotated	Coverage RES1 (rpkm)[c]	Coverage RES2 (rpkm)[c]
c2	*Bacteroidetes Cytophagia Cytophagale Flammeovirgaceae*	*Marivirga tractuosa* (50.15% AAI)	91.9	2.7	78.4	41.79	396	3.87	82 589	14 095	3747	1452 (38.8%)	0.1 ± 0.07	0.4 ± 0.46
c3	*Alphaproteobacteria Rhodospirillales Acetobacteraceae*	*Roseomonas* sp. (56.54% AAI)	94.6	0.9	90.1	69.05	292	4.52	125 874	24 612	4612	2024 (43.9%)	1.32 ± 2.46	1.35 ± 1.71
c4	*Alphaproteobacteria Caulobacterales Caulobacteraceae*	*Caulobacteraceae* bacterium (52.29% AAI)	91	0	91	67.51	163	2.72	83 313	25 390	2817	1389 (49.3%)	0.79 ± 1.16	1.63 ± 1.3
c14	*Gammaprotebacteria*	*Thiohalobacter thiocyanaticu*s (48.24% AAI)	90.1	1.8	81.1	67.88	729	5.51	34 030	9 353	5556	2325 (41.8%)	0.14 ± 0.1	0.56 ± 0.53
c19	*Alphaproteobacteria Sphingomonadales Erythrobacteraceae Porphyrobacter*	*Porphyrobacter* sp. CACIAM (81.5 ± 14,4% AAI)	94.6	0.9	90.1	64.06	58	3.41	476 516	264 474	3279	1586 (48.4%)	0.05 ± 0.01	4.6 ± 7.23
c24	*Betaproteobacteria Burkholderiales*	*Rhizobacter gummiphilus* (62.69% AAI)	95.5	1.8	86.5	64.87	64	3.46	232 861	99 668	3260	1832 (56.2%)	0.09 ± 0.06	0.49 ± 0.53
c26	*Betaproteobacteria Methylophilales Methylophilaceae Methylotenera*	*Methylotenera mobilis* (62.03% AAI)	95.5	1.8	86.5	41.94	12	2.03	735 558	469 317	2016	1268 (62.9%)	0.02 ± 0.01	0.58 ± 1.17
c27	*Alphaproteobacteria Rhizobiales Bradyrhizobiaceae*	*Bosea* sp. PAMC (79.03 ± 15.87%	95.5	0.9	91	66.75	117	4.81	204 440	72 057	4670	2376 (50.9%)	0.06 ± 0.02	2.6 ± 3.8

	Bosea	AAI)												
c31	*Alphaproteobacteria Rhizobiales Beijerinckiaceae*	*Methylocella silvestris* (56.71% AAI)	95.5	1.8	86.5	62.55	58	4.43	399 037	106 052	4266	1942 (45.5%)	0.08 ± 0.06	0.78 ± 1.06
c33	*Alphaproteobacteria Rhizobiales Bradyrhizolbiaceae Oligotropha*	*Oligotropha carboxidovorans* (83.31 ± 12.94% AAI)	94.6	1.8	85.6	61.68	51	3.45	253 596	122 000	3301	1794 (54.3%)	0.05 ± 0.03	0.55 ± 0.82
c34	*Gemmatimonadetes Gemmatimonadales Gemmatimonadaceae Gemmatirosa*	*Gemmatimonas phototrophica* (53.28% AAI)	73	1.8	64	67.83	537	2.74	26 761	5 888	2892	1226 (42.4%)	0.16 ± 0.08	0.51 ± 0.47
c35	*Alphaproteobacteria Rhizobiales Hyphomicrobiaceae Hyphomicrobium*	*Hyphomicrobium denitrificans* (87.12 ±14.44% AAI)	88.3	1.8	79.3	58.89	431	4.00	85 476	13 110	4239	1835 (43.3%)	2.3 ± 2.22	1.12 ± 1
c38	*Alphaproteobacteria Sphingomonadales Sphingomonadaceae Novosphingobium*	*Novosphingobium aromaticivorans* (66.34% AAI)	93.7	0.9	89.2	68.83	64	4.07	251 248	119 692	3727	1770 (47.5%)	0.07 ± 0.07	0.53 ± 0.54
c39	*Betaproteobacteria Sulfuricellales Sulfuricellaceae Sulfuricella*	*Sulfuricella denitrificans* (84.31 ± 17.04% AAI)	94.6	3.6	76.6	58.61	92	3.57	180 696	66 989	3485	1851 (53.1%)	0.07 ± 0.02	0.5 ± 1.02
c41	*Alphaproteobacteria Sphingomonadales Erythrobacteraceae Porphyrobacter*	*Porphyrobacter* sp. LM 6 (97.28 ± 7.93% ANI)	93.7	1.8	84.7	64.72	33	3.04	397 646	156 172	2951	1471 (49.8%)	0.17 ± 0.1	3.74 ± 8.4
c43	*Alphaproteobacteria Rhodospirillales*	*Azospirillum brasilense* (49.02% AAI)	93.7	2.7	80.2	66.19	611	5.89	51 990	13 561	6212	2683 (43.2%)	0.5 ± 0.56	0.38 ± 0.38

c47	*Betaproteobacteria Nitrosomonadales Gallionellaceae Sideroxydans*	*Sideroxydans lithotrophicus* (77.27 ± 16,96% AAI)	73	2.7	59.5	55.31	455	2.29	26 046	5 807	2655	1512 (56.9%)	0.3 ± 0.35	0.21 ± 0.21
c48.1 [a]	*Alphaproteobacteria Sphingomonadales Sphingomonadaceae Sphingopyxis*	*Sphingopyxis terrae* (63.28% AAI)	94.6	1.8	85.6	63.12	29	2.95	433 016	176 469	2853	1429 (50.1%)	0.06 ± 0.03	1.85 ± 3.71
c48.2 [a]	*Alphaproteobacteria Rhizobiales Bradyrhizobiaceae*	*Bradyrhizobium lablabi* (64.55% AAI)	94.6	0.9	90.1	62.2	16	4.57	902 226	479 955	4332	1996 (46.1%)	0.04 ± 0.01	1.68 ± 2.22
c49	*Alphaproteobacteria Rhizobiales*	*Rhizobium tropici* (45.25% AAI)	87.4	0.9	82.9	69.43	978	4.80	38 978	5 604	5423	2260 (41.7%)	0.23 ± 0.28	0.6 ± 0.68
c51	*Nitrospira Nitrospirales Nitrospiraceae Nitrospira*	*Nitrospira moscoviensis* (62.69% AAI)	92.8	5.4	65.8	57.78	80	4.27	1 125 973	247 953	4354	1608 (36.9%)	16.93 ± 20.05	4.36 ± 6.92
c56	*Nitrospira Nitrospirales Nitrospiraceae Nitrospira*	*Nitrospira moscoviensis* (62.1% AAI)	77.5	1.8	68.5	57.14	443	3.69	44 036	11 519	4121	1498 (36.4%)	0.43 ± 0.5	0.23 ± 0.17
c58	*Betaproteobacteria Nitrosomonadales Nitrosomonadaceae Nitrosomonas*	*Nitrosomonas* sp. Is79A3 (66.34% AAI)	94.6	0.9	90.1	48.69	70	3.12	237 347	90 243	2951	1486 (50.4%)	7.12 ± 10.39	22.12 ± 27.3
c59	*Alphaproteobacteria Rhodobacterales Rhodobacteraceae*	*Rhodobacter* sp. CZR27 (62.45% AAI)	93.7	1.8	84.7	66.06	359	4.33	103 413	17 991	4620	2192 (47.4%)	0.14 ± 0.08	0.51 ± 0.85
c60	*Betaproteobacteria Nitrosomonadales Methylophilaceae*	*Methylotenera mobilis* (77.35 ± 15,15%	95.5	1.8	86.5	45.61	24	2.53	486 922	217 332	2401	1424 (59.3%)	0.02 ± 0.01	2.44 ± 4.84

	Methyloteneva	AAI)										
c61	*Betaproteobacteria Nitrosomonadales Gallionellaceae Sideroxydans/Gallionella*	*Sideroxydans lithotrophicus* (67.0% AAI)	94.6	0.9	90.1	56.92	50	2.75	341 276	100 797	2814	1477 (52.5%)	1.65 ± 1.82	0.09 ± 0.1
c64	*Alphaproteobacteria Caulobacterales*	*Caulobacteraceae* bacterium (52.17% AAI)	87.4	0	87.4	65.98	313	2.43	54 901	10 673	2708	1354 (50.0%)	0.8 ± 0.87	0.27 ± 0.26
c65	*Alphaproteobacteria Sphingomonadales Sphingomonadaceae*	*Sphingomonas panacis* (61.83% AAI)	93.7	0	93.7	61.44	33	3.23	571 943	183 848	3132	1541 (49.2%)	0.99 ± 1.73	1.18 ± 2.5
c69	*Betaproteobacteria Nitrosomonadales Nitrosomonadaceae*	*Nitrosospira lacus* (52.37% AAI)	93.7	1.8	84.7	60.83	82	3.91	228 320	102 579	3874	1967 (50.8%)	0.05 ± 0.03	2.75 ± 4.57
c70	*Alphaproteobacteria Sphingomonadales Sphingomonadaceae Sphingomonas*	*Sphingomonas panacis* (61.38% AAI)	94.6	3.6	76.6	63.51	44	3.14	382 616	247 649	3082	1600 (51.9%)	10.74 ± 18.11	11.95 ± 17.62
c72	*Alphaproteobacteria Sphingomonadales Sphingomonadaceae Sphingomonas*	*Sphingomonas panacis* (62.63% AAI)	93.7	0	93.7	61.81	31	3.89	611 207	391 571	3818	1682 (44.1%)	2.76 ± 3.5	2.84 ± 5.13
c74	*Alphaproteobacteria Rhizobiales Hyphomicrobiaceae*	*Hyphomicrobium nitrativorans* (54.07% AAI)	92.8	0.9	88.3	65.43	103	5.94	294 341	93 246	5453	2251 (41.3%)	0.33 ± 0.72	1.07 ± 1.41
c76	*Betaproteobacteria Nitrosomonadales Gallionellaceae Gallionella*	*Gallionella capsiferriformans* (79.91 ± 16.11% AAI)	92.8	0.9	88.3	53.15	130	3.13	121 681	37 955	3102	1586 (51.1%)	0.05 ± 0.02	0.34 ± 0.69
c77	*Betaproteobacteria Nitrosomonadales*	*Methylotenera versatilis*	95.5	2.7	82	42.38	11	2.36	828 080	282 428	2257	1324 (58.7%)	0.03 ± 0.03	0.62 ± 1.76

	Methylophilaceae *Methylotenera*	(64.03% AAI)											
c83	*Betaproteobacteria Nitrosomonadales Gallionellaceae Sideroxydans/Gallionella*	*Gallionella capsiferriformans* (62.29% AAI)	95.5	0.9	91	56.01	132	3.01	195 785	37 811	2977	1581 (53.1%)	0.05 ± 0.03	0.6 ± 1.5
c86	*Alphaproteobacteria Rhizobiales*	*Rhizobium* sp. N324 (47.54% AAI)	81.1	5.4	54.1	63.81	669	3.19	45 151	5 479	3618	1658 (45.8%)	0.2 ± 0.23	0.52 ± 0.6
c90	*Alphaproteobacteria Rhizobiales Hyphomicrobiaceae*	*Hyphomicrobium nitrativorans* (54.02% AAI)	93.7	0.9	89.2	63.9	87	5.15	412 933	130 872	4854	2111 (43.5%)	4.07 ± 2.22	2.77 ± 0.97
c93	*Alphaproteobacteria Rhizobiales Methylobacteriaceae Methylobacterium*	*Methylobacterium* sp. C1 (65.19% AAI)	94.6	3.6	76.6	66.32	231	5.46	316 841	51 541	5451	2106 (38.6%)	1.93 ± 3.36	2.57 ± 4.3
c94	*Planctomycetes Planctomycetia Planctomycetales Planctomycetaceae*	*Singulisphaera acidiphila* (41.76% AAI)	91.9	2.7	78.4	46.18	124	4.80	525 508	101 160	3712	1329 (35.9%)	0.12 ± 0.08	0.44 ± 1.02
c97	Unknown *Planctomycetes Phycisphaerae*	*Phycisphaera mikurensis* (40.47% AAI)	89.2	0	89.2	65.03	76	3.60	258 711	82 719	2976	1133 (38.1%)	0.03 ± 0.02	0.68 ± 0.92
c102	*Betaproteobacteria Nitrosomonadales Gallionellaceae Sideroxydans/Gallionella*	*Sideroxydans lithotrophicus* (66.35% AAI)	93.7	1.8	84.7	55.38	116	2.63	140 207	39 153	2675	1520 (56.8%)	0.16 ± 0.07	0.23 ± 0.42
c103.1[a]	*Alphaproteobacteria Rhizobiales Hyphomicrobiaceae*	*Hyphomicrobium denitrificans* (61.59% AAI)	95.5	7.2	59.5	59.99	12	3.65	1 305 095	829 204	3405	1705 (50.1%)	1.92 ± 1.4	2.01 ± 1.75

	Hyphomicrobium													
c103. 2[a]	*Alphaproteobacteria Rhizobiales*	*Chelatococcus* sp. (51.55% AAI)	94.6	0.9	90.1	60.73	67	3.28	248 172	102 576	3179	1765 (55.5%)	24.54 ± 12.15	19.91 ± 10.14
c104	*Alphaproteobacteria Sphingomonadales Sphingomonadaceae*	*Sphingomonas panacis* (61.76% AAI)	94.6	0.9	90.1	66.91	96	3.25	161 741	54 966	3263	1554 (47.6%)	4.91 ± 6.11	3.59 ± 2.98
c107	*Betaproteobacteria Nitrosomonadales Nitrosomonadaceae Nitrosomonas*	*Nitrosomonas* sp. Is79A3 (66.17% AAI)	93.7	4.5	71.2	48.3	163	3.66	208 136	64 726	3461	1687 (48.7%)	42.54 ± 26.31	11.82 ± 6.23
c109	*Betaproteobacteria Nitrosomonadales Nitrosomonadaceae*	*Nitrosospira* lacus (52.47% AAI)	95.5	0.9	91	62.68	59	3.80	299 525	118 187	3636	1832 (50.4%)	0.33 ± 0.27	1.78 ± 4.76
c114	*Betaproteobacteria Burkholderiales Comamonadaceae*	*Acidovorax* sp. NA3 (65.46% AAI)	91.9	2.7	78.4	63.89	497	3.99	114 432	12 659	4261	2171 (51.0%)	0.14 ± 0.06	0.91 ± 2.04

[a] Manually curated Metagenome Assembled Genomes (MAGs)

[b] Taxon was assigned when at least 75% of the identified genes resulted in a concordant taxonomy.

[c] Values calculated as the mean coverage of each MAG across all sample within each reservoir (Mean ± Standard deviation)

Table C7: Percentage coverage of SSU rRNA contigs identified as bacterial phyla, eukaryota and unclassified averaged across both reservoirs (coverage calculated for contigs > 250bp)

			RES1		RES2	
			MRA	SD	MRA	SD
Bacterial phyla	*Proteobacteria*		81.60	13.32	82.98	10.28
		Alphaproteobacteria	42.99	13.97	52.75	15.36
		Deltaproteobacteria	0.08	0.08	0.09	0.09
		Total *Gammaproteobacteria*	38.53	14.54	30.15	17.02
		Gammaproteobacteria, Betaproteobacteriales	38.05	14.47	29.46	17.14
		Other *Gammaproteobacteria*	0.48	0.37	0.69	0.35
	Nitrospirae		11.25	13.48	2.28	3.83
	Bacteroidetes		0.40	0.33	0.80	0.88
	Planctomycetes		0.16	0.14	0.58	0.94
	Actinobacteria		0.43	0.26	0.30	0.29
	Acidobacteria		0.12	0.12	0.27	0.28
	Gemmatimonadetes		0.19	0.18	0.10	0.16
	Patescibacteria		0.11	0.10	0.08	0.10
	Verrucomicrobia		0.04	0.07	0.10	0.10
	Cyanobacteria		0.02	0.03	0.09	0.11
	Spirochaetes		0.03	0.06	0.06	0.11
	Chloroflexi		0.03	0.07	0.05	0.06
	Chlamydiae		0.03	0.05	0.02	0.03
Eukaryota			2.05	1.26	4.64	5.55
Unclassified contigs			3.55	1.17	7.58	3.66

Table C8: Percentage coverage of SSU rRNA contigs identified as Eukaryota phyla averaged across both reservoirs (coverage calculated for contigs > 250bp)

	RES1		RES2	
	MRA	**SD**	**MRA**	**SD**
Metazoa	0.20	0.14	1.12	1.79
Unclassified Eukaryota	0.31	0.08	0.92	1.24
Opisthokonta	0.61	1.29	0.49	1.01
Ochrophyta	0.67	0.65	0.39	0.69
Stramenopiles	0.01	0.00	0.16	0.33
Amoebozoa	0.03	0.09	0.06	0.11
Bacillariophyta	0.01	0.02	0.05	0.04
Viridiplantae	0.03	0.08	0.04	0.07
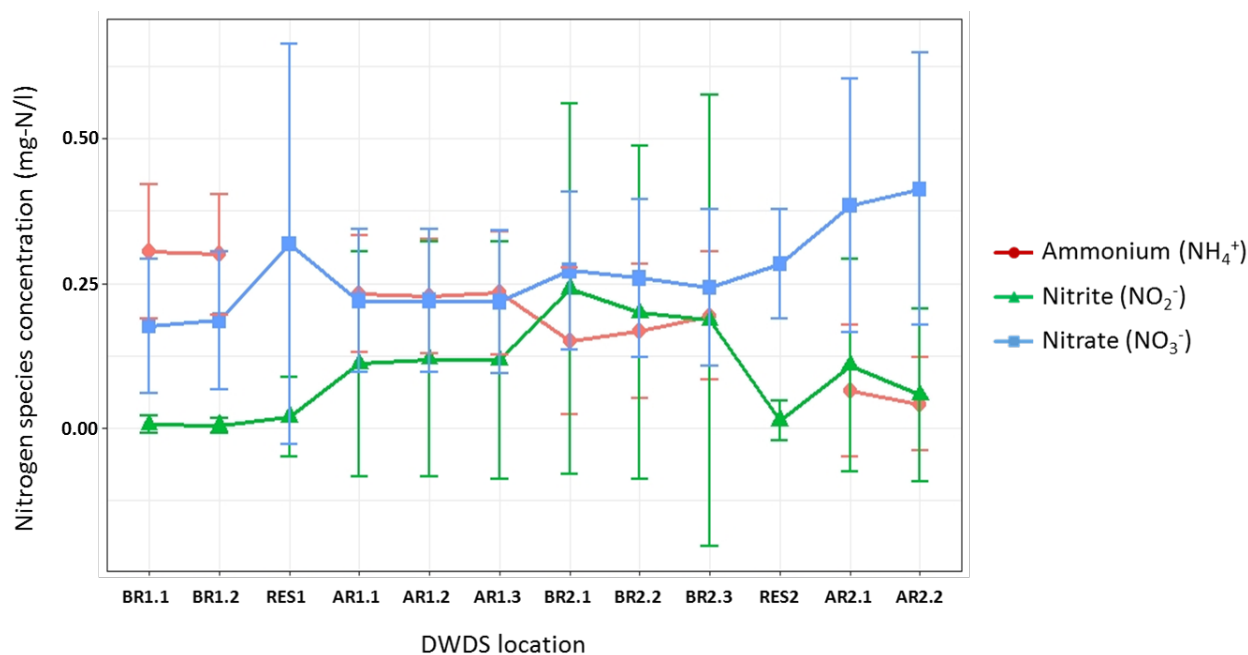
Figure C1: Spatial changes in the concentrations of nitrogen species [i.e., ammonium (NH_4^+), nitrite (NO_2^-) and nitrate (NO_3^-)] for locations within the chloraminated section of the DWDS averaged over two years. Error bars capture variation in concentration of each nitrogen species over the period of two years. Location names on the x-axis translate to: before reservoir 1 (BR1.1 and BR1.2), reservoir 1 (RES1), after reservoir 1 (AR1.1 – AR1.3), before reservoir 2 (BR2.1 – BR2.3), reservoir 2 (RES2) and locations after reservoir 2 (AR2.1 and AR2.2). Ammonium concentration data is missing for sites RES1 and RES2.
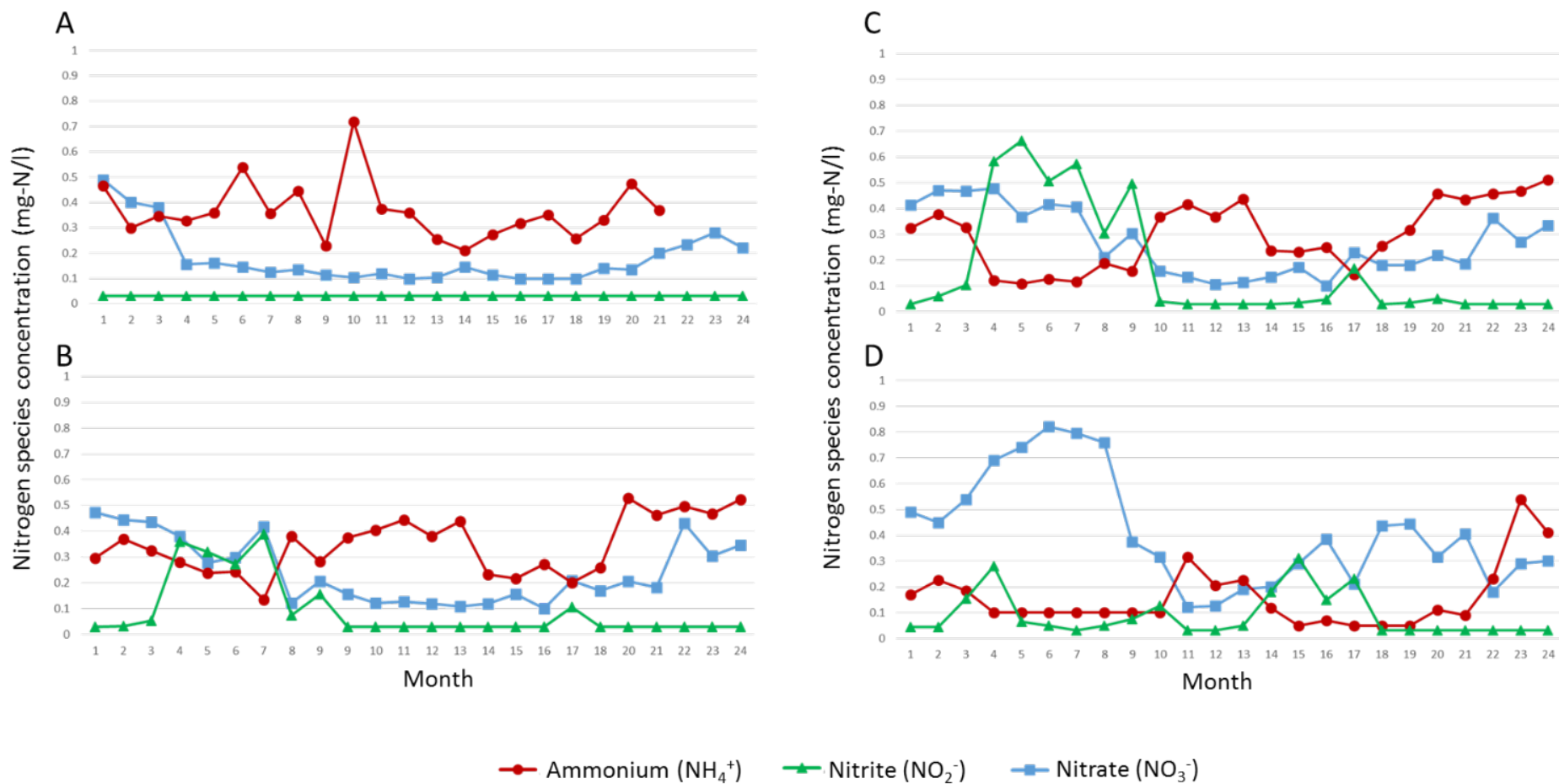
Figure C2: Temporal changes in nitrogen compounds [i.e., ammonium (red circles), nitrite (green triangles) and nitrate (blue squares)] over 2 years. Plots (A) and (B) refer to before and after RES1, respectively and plots (C) and (D) refer to before and after RES2, respectively.
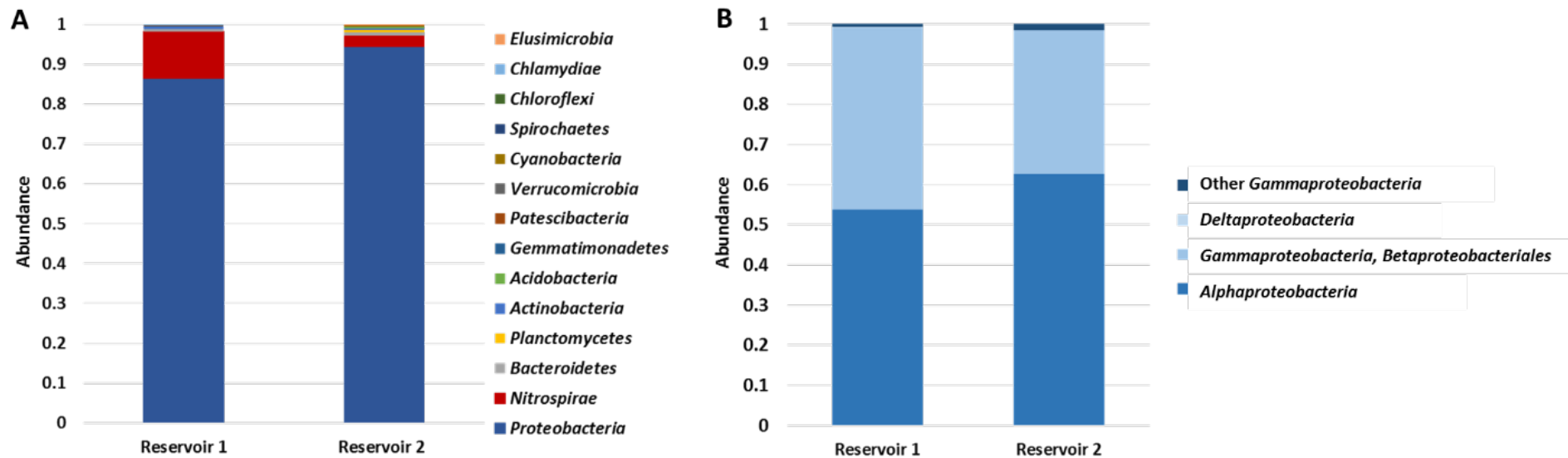
Figure C3: The relative abundance of bacterial SSU rRNA genes i.e., bacterial phyla (A) and proteobacterial classes (B). The relative abundance of SSU rRNA genes within each taxonomic grouping was averaged across all samples for each reservoir (RES1 and RES2) (Coverage calculated for contigs > 250bp).
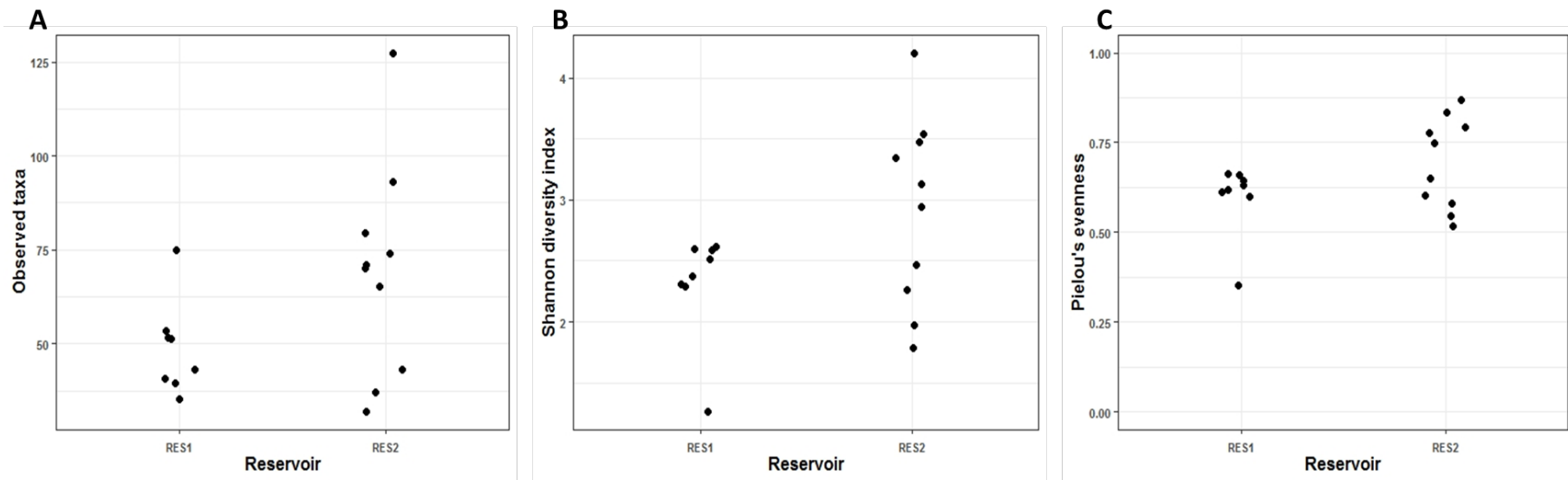
Figure C4: Alpha diversity measures of samples from both reservoirs (RES1 and RES2) i.e. richness (A), Shannon diversity (B) and evenness (C).
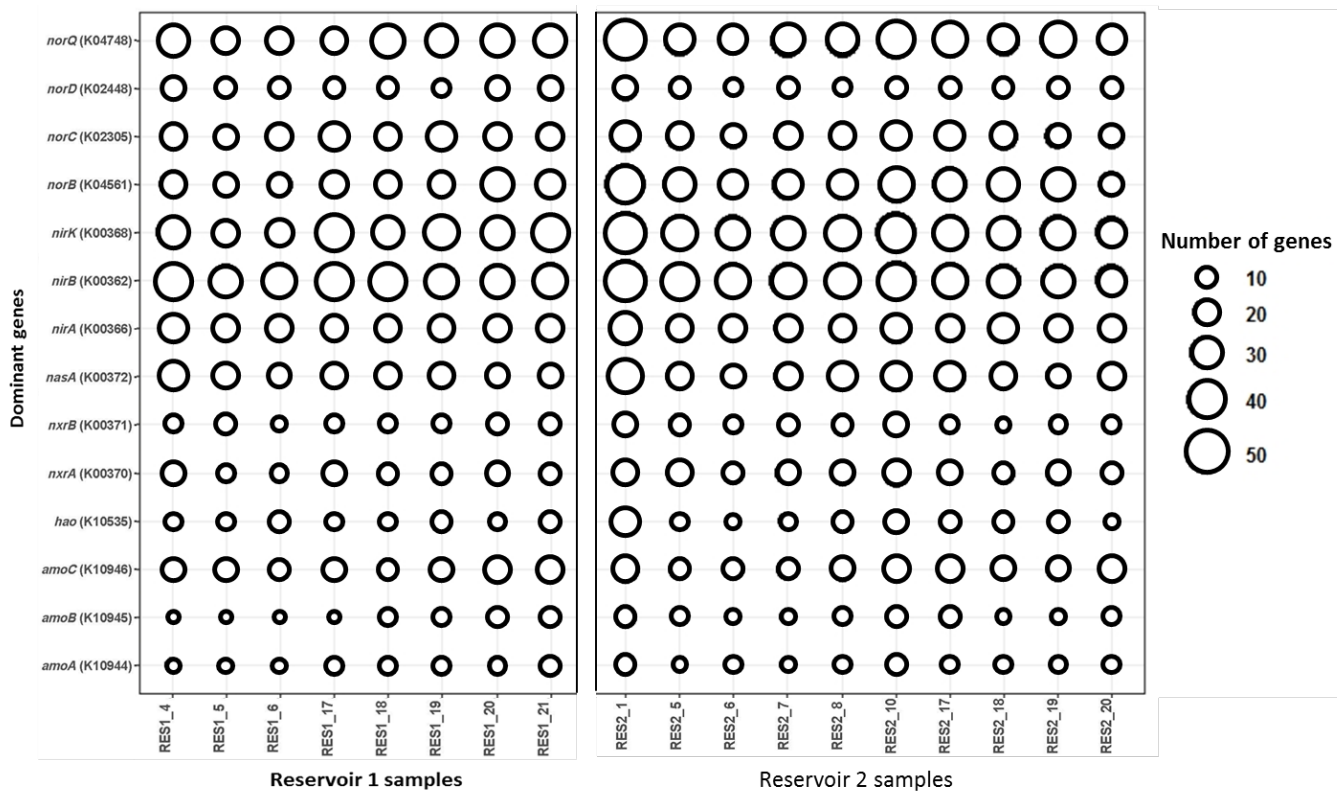
Figure C5: The number of genes identified for each dominant function across all reservoir samples.
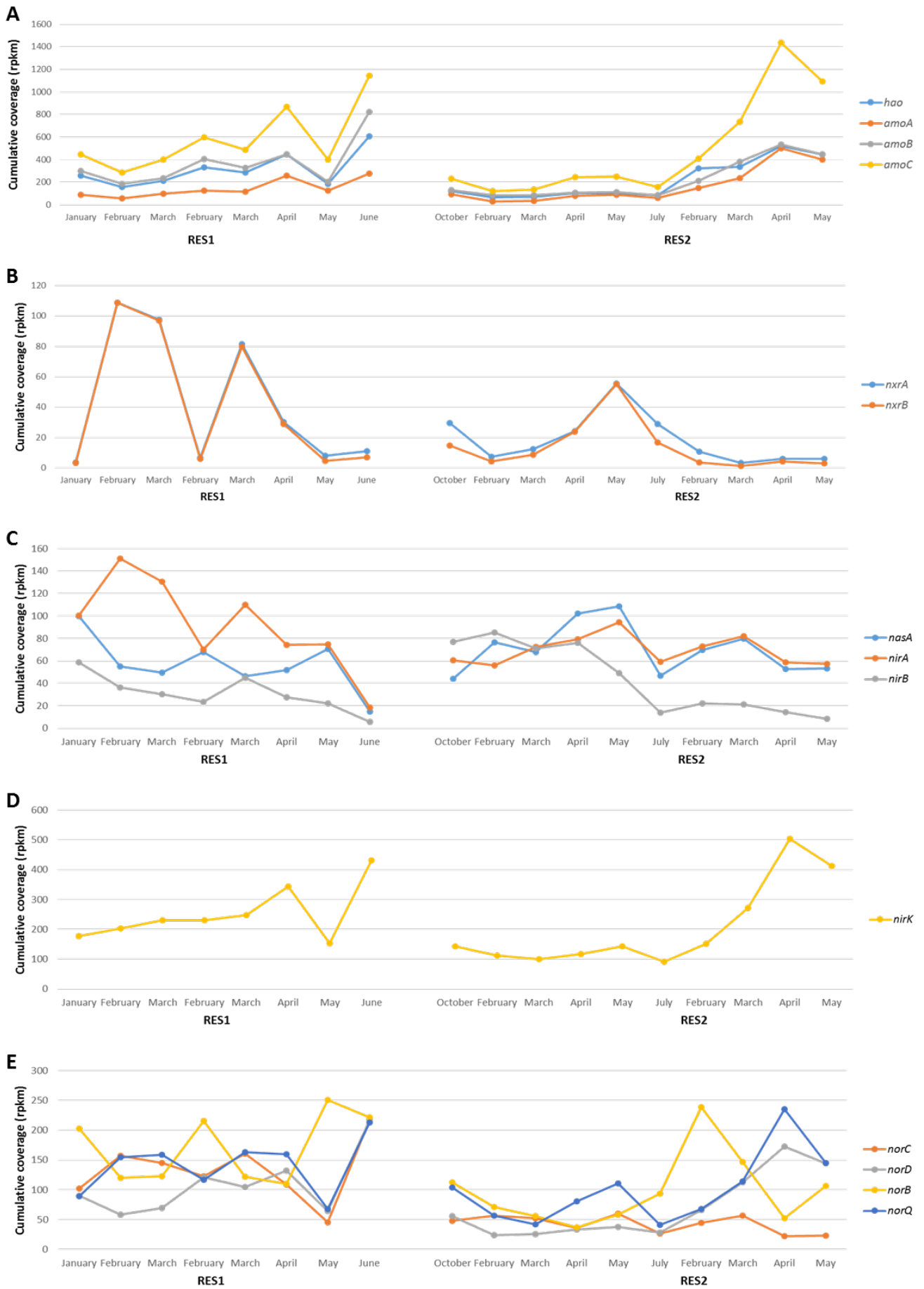
Figure C6: Temporal trends in the cumulative coverage of the dominant nitrogen transforming genes identified in the dataset. (A) ammonia oxidation (*amoABC* and *hao* genes), (B) nitrite oxidation (*nxrAB* genes), (C) assimilatory nitriate and nitrite reduction (*nasA* and *nirAB* genes), (D) nitrite oxidation, nitric oxide-forming (*nirK* gene) and (E) nitric oxide reduction (*norBCDQ* genes).
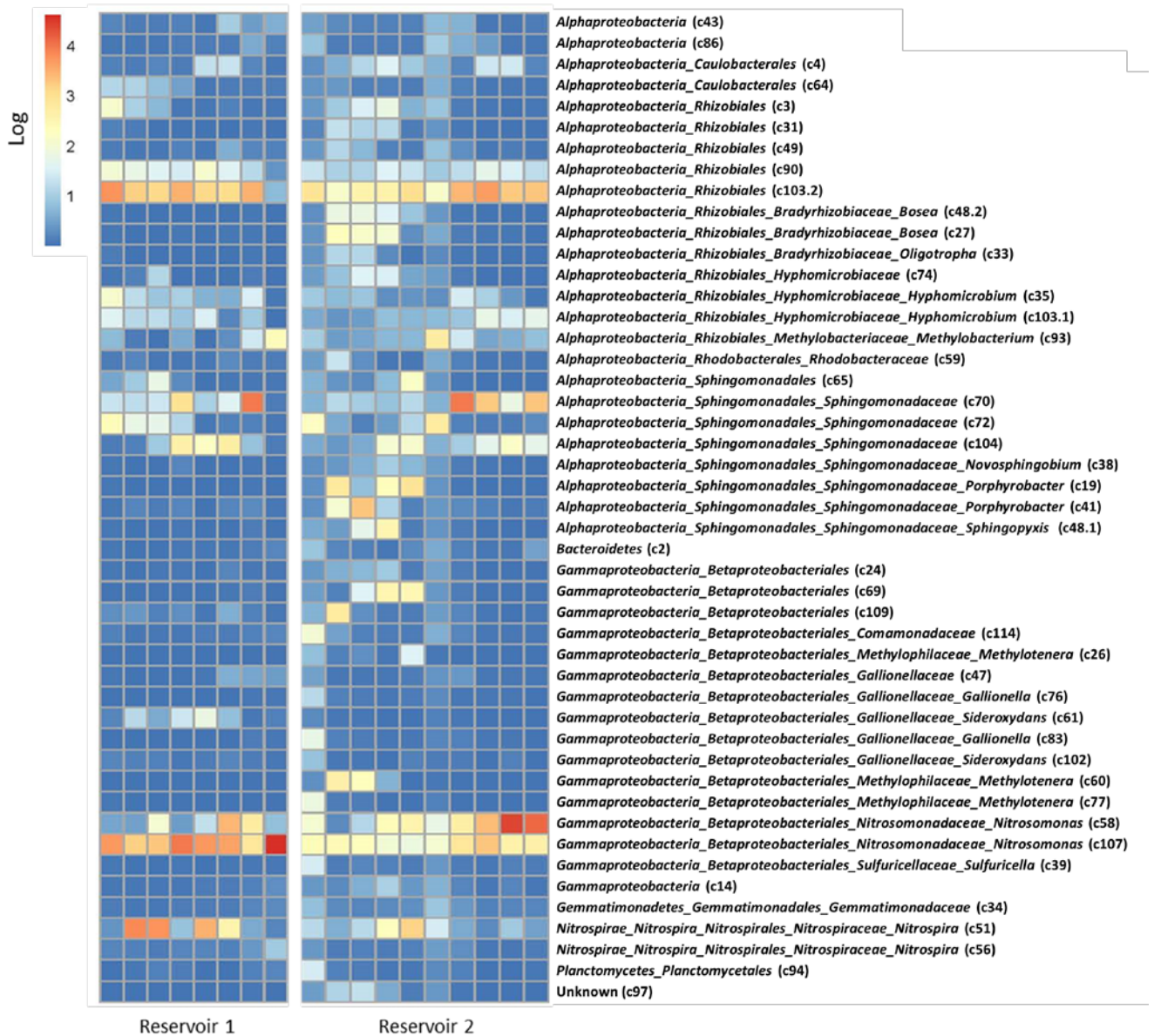
Figure C7: Log transformed coverage of the 47 Metagenome Assembled Genomes (MAGs) across samples from the two reservoirs. The level of coverage is indicated by the key on the top left of the figure as the log of reads per kilo base million (rpkm).
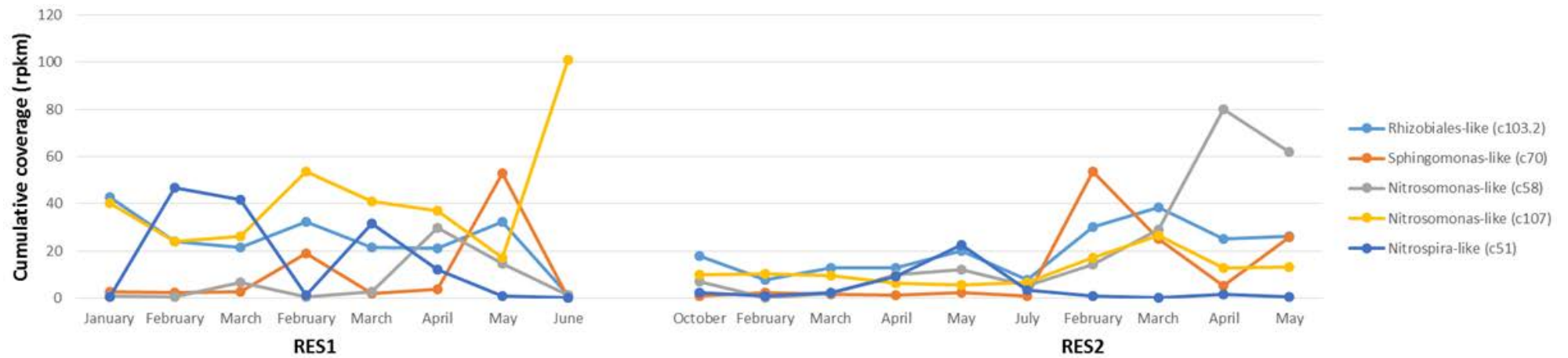
Figure C8: Temporal trends in the cumulative coverage of the dominant Metagenome Assembled Genomes (MAGs) constructed in the dataset.
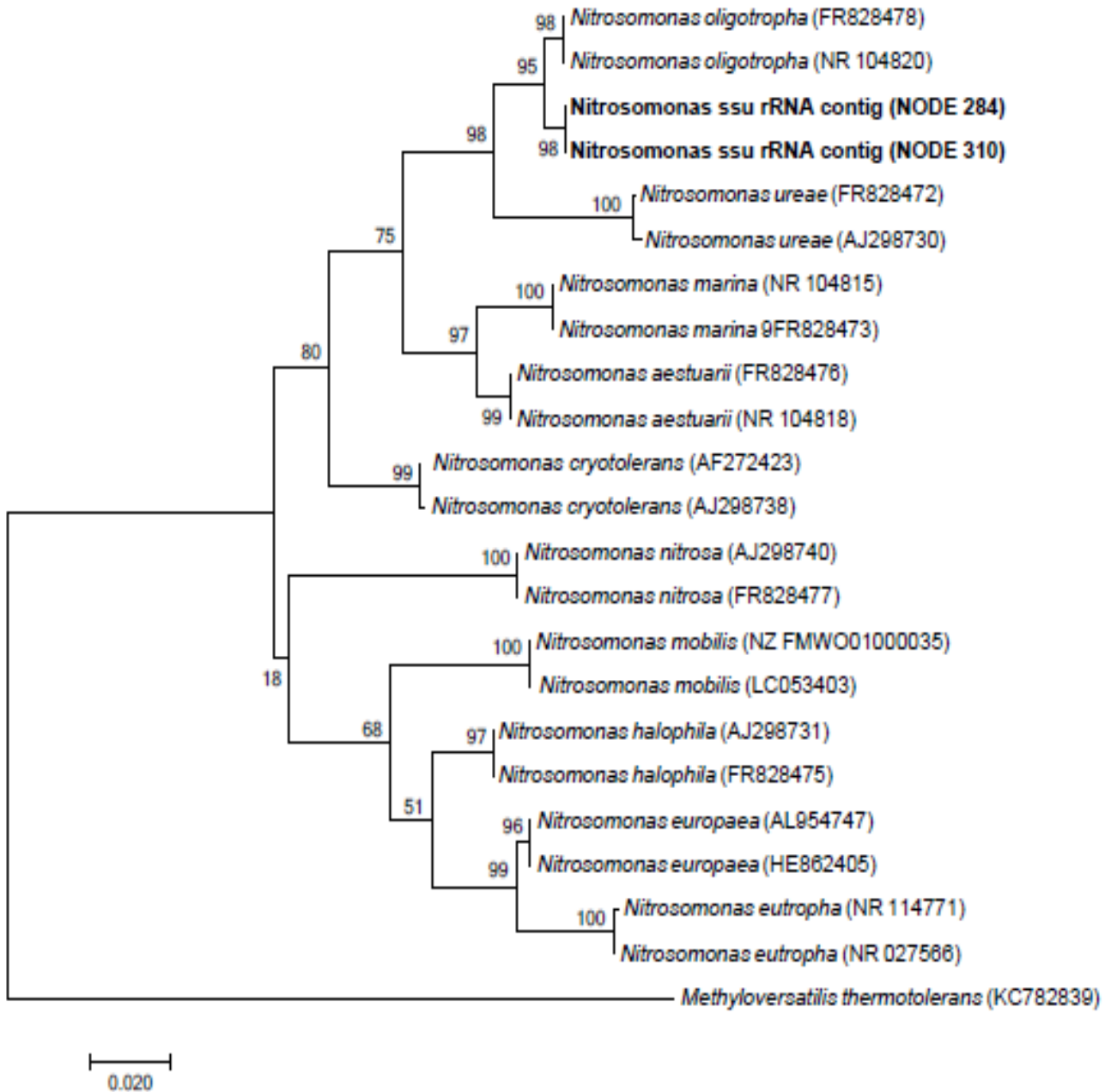
Figure C9: Maximum likelihood phylogenetic tree showing the grouping of *Nitrosomonas* SSU rRNA contigs retrieved from metagenomic data with *Nitrosomonas* spp. 16S rRNA reference strains. With *Methyloversatilis thermotolerans* as the outgroup and bootstrap analysis of 1000 replicates (bootstrap values are indicated as percentages).
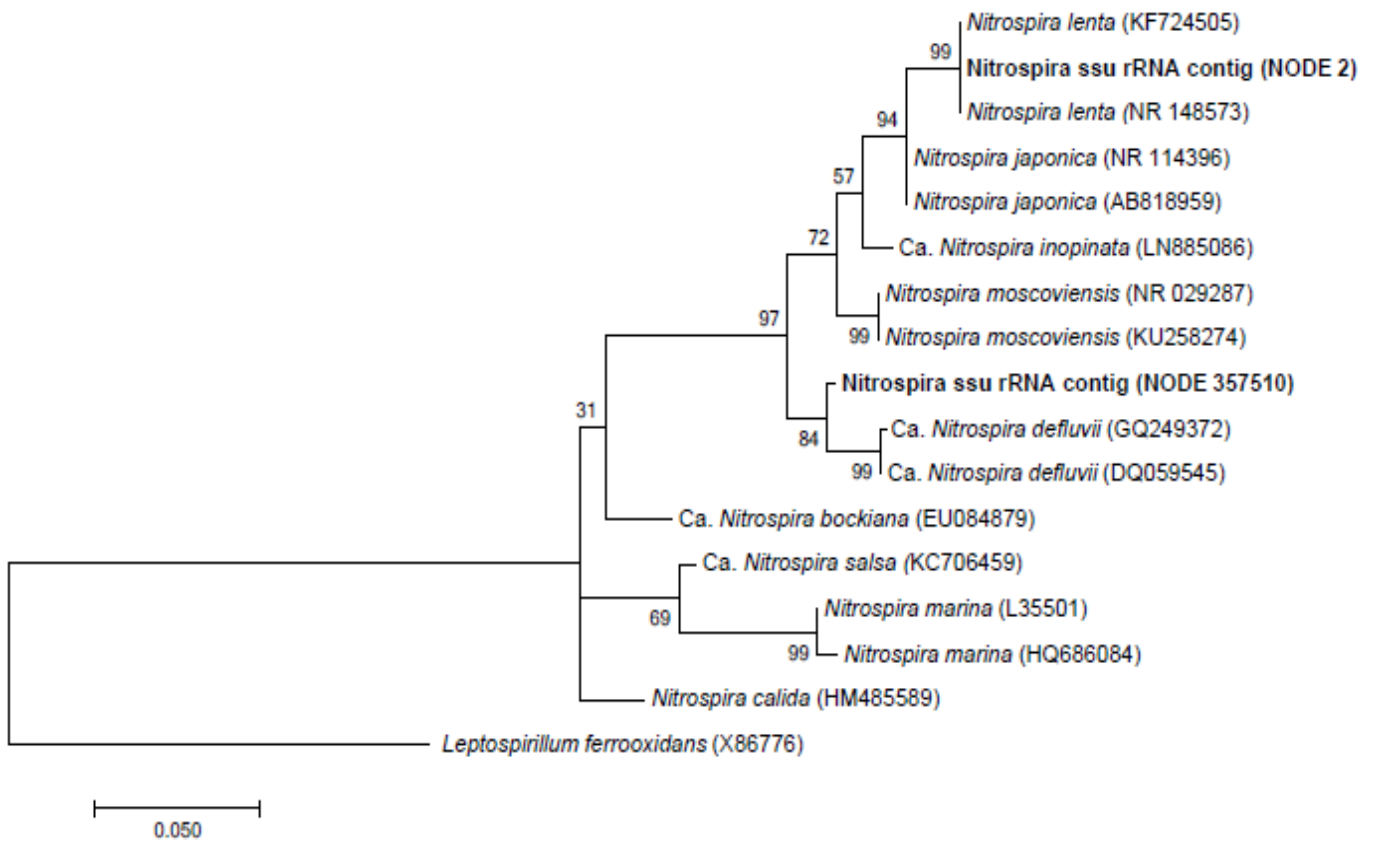
Figure C10: Maximum likelihood phylogenetic tree showing the grouping of *Nitrospira* SSU rRNA contigs retrieved from metagenomic data with *Nitrospira* spp. 16S rRNA reference strains. With *Leptospirillum ferrooxidans* as the outgroup and bootstrap analysis of 1000 replicates (bootstrap values are indicated as percentages).
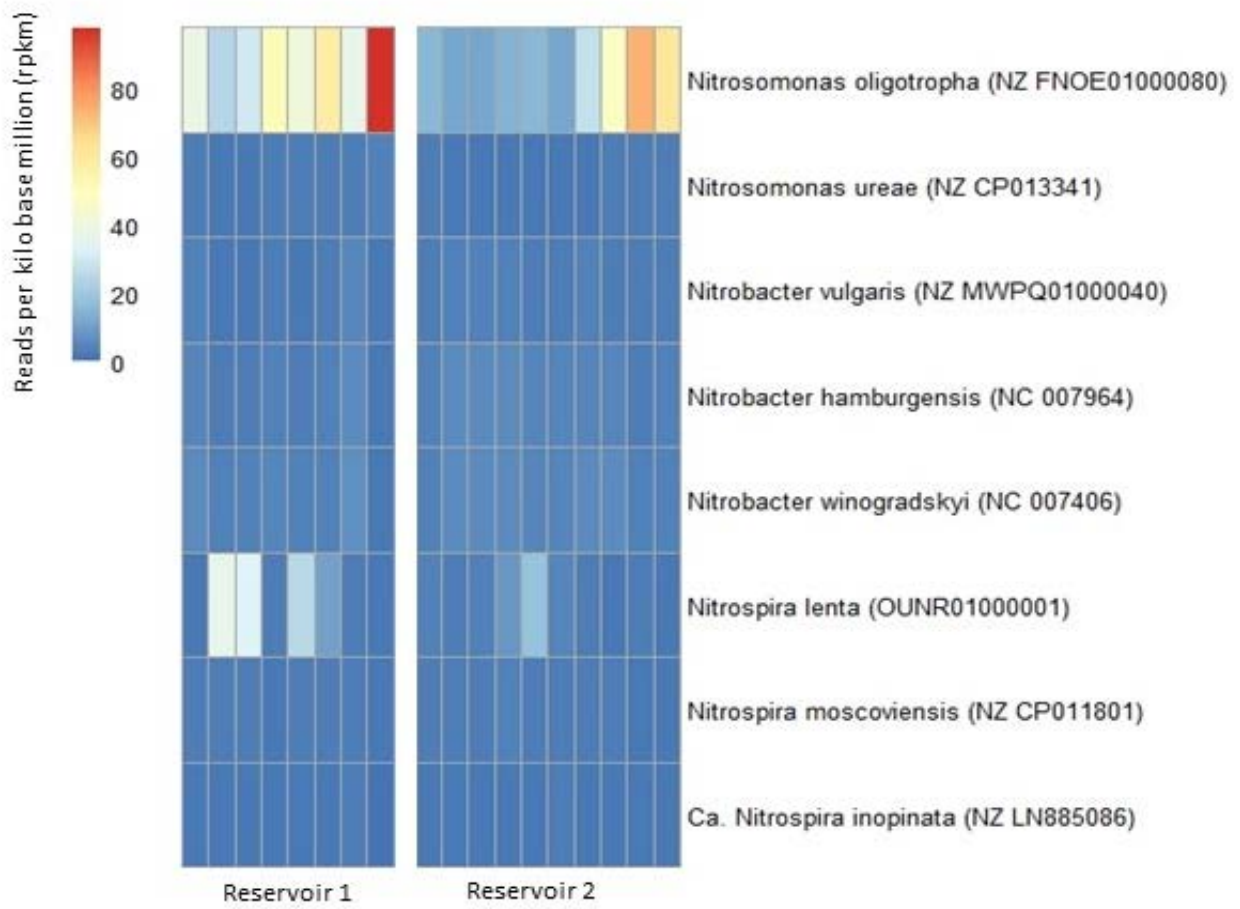
Figure C11: Heatmap showing the coverage of reference genomes across all samples, which mapped on average > 1 million reads per kilo base of the reference genome (rpkm).

CHAPTER 5

CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS

From the literature it is clear that individual drinking water systems differ significantly from each other. These differences are attributed to the chemical, physical and biological factors inherent to each drinking water system (i.e. in treatment and distribution), which ultimately shape the drinking water microbiome and drive microbial community dynamics. Drinking water systems are multi-dimensional complex systems, where these factors have an overall combined effect on the microbial community and ultimately water quality. The complexity of interactions between the contributing factors, differing source waters, the variability between distribution systems and the difficulty in standardisation of those systems makes it difficult to predict the microbial community dynamics across different drinking water systems.

Therefore, the overall goal of this study was to comprehensively understand the microbial community dynamics within a large-scale South Africa drinking water treatment and distribution system (on various spatial and temporal scales), while considering the impact of various chemical, physical and biological factors on the drinking water microbiome. This study also attempted to investigate the reproducibility of the observed microbial community dynamics in different sections of the DWDS. And lastly, due to the use of chloramine as a secondary disinfectant, the dominance of *Nitrosomonas* spp. and the increased potential for nitrification, this study aimed to determine the genetic potential of the microbial community behind nitrogen metabolism.

Initially, this study involved a long-term sampling campaign of two years, where the temporal and spatial dynamics of the microbial community were characterised, using 16S rRNA gene profiling. Focusing on the DWDS, monthly bulk water samples were collected from the outlet of the treatment plant and 17 locations along the distribution system, covering multiple disinfectant strategies. Through this long-term, high-frequency investigation, the seasonal and annual patterns in the microbial community were observed that would have otherwise been missed in a short-term study. As a result, strong temporal trends and seasonal cycling were observed in community richness, diversity and structure, which correlated to changes in water temperatures and disinfectant residual concentrations. It was also observed that temporal trends dominated on a localised level, at individual sample locations, but became less evident as bulk water moved away from the DWTP. Spatial dynamics revealed distance decay features, where the microbial community became increasingly dissimilar with increasing distance between samples locations. Overall, when considering the DWDS in its entirety, spatial dynamics of the DWDS were stronger and explained

more of the variation in the microbial community composition and structure. This may be attributed to the different disinfectant residuals applied in this system.

The drinking water system under investigation in this study was unique in the sense that differing source waters are abstracted from a dam and river system and consequently treated at two separate DWTPs. The same consecutive treatment strategies were applied at each DWTP. Chlorinated water leaving the two DWTPs then supplied different sections within the same large-scale DWDS. Therefore, this system provided a unique opportunity to study the reproducibility of the microbial community dynamics of the two drinking water systems (i.e. the DWTP and their corresponding DWDS sections) treating similar source waters. Here, corresponding samples from the two DWTPs and DWDSs were obtained for 8 months. Using 16S rRNA gene profiling, reproducible microbial dynamics were observed in both DWTPs and their corresponding DWDS sections. Although, source waters showed differences in microbial community structure and membership, the treatment processes (coagulation, flocculation, sedimentation and carbonation) in both DWTPs caused increased community similarity and selected for the same dominant taxa. This suggested that similarities in operational parameters and design of the two DWTP resulted in the development of similar microbial communities. Conversely, chlorination resulted in increased dissimilarity and temporal variation between the two systems, indicating differing responses of the microbial community to disinfection. Following disinfection the impact of distribution was observed between corresponding samples from the two DWDS. Here it was observed that certain dominant taxa were selected for, indicating stabilisation of the microbial community post-disinfection. Although, dissimilarities inherent to each system were observed, treatment and distribution had a similar impact on the microbial community dynamics. The observed differences between the two systems may be attributed to site specific rare/low abundant taxa as a result of the initial differences in the source waters and the differential response to chlorination.

Thus far, these studies have demonstrated that a 16S rRNA gene community profiling approach provided valuable information regarding the effects of treatment and distribution on the spatial and temporal dynamics of the drinking water microbial community composition and structure. However, studies linking microbial community composition and diversity to function are limited. Here, metagenomics provides an introduction into the genetic potential of a community's functional and metabolic activity, at a specific time point. Using this approach changes in response to environmental disturbances or existing conditions can be examined.

In both the previous studies, the observed dominance of *Nitrosomonas* spp. in the chloraminated sections of the DWDS suggested that ammonia oxidation and ultimately nitrification may be important processes in this DWDS. Here, the addition of ammonia through chloramination provided an alternative source of nitrogen in the system. With this in mind, the metabolic fate of nitrogen in the chloraminated drinking water was investigated. Using a shotgun metagenomic approach, the understanding of nitrogen cycling potential of the microbial community was improved in chloraminated drinking water. Select bulk water samples were collected from two reservoirs within the chloraminated section of the DWDS, within the 2 year sampling campaign. This study revealed the genes and bacteria responsible for the metabolism of nitrogen and other nitrogen transforming reactions. Specifically, potential nitrification was observed as nitrate concentrations increased as the distance from the point of chloramination increased. Here, the observed changes in the concentration of nitrogen species and the dominance of *Nitrosomonas* and *Nitrospira*-like bacteria suggested that they play a significant role in contributing to varying stages of nitrification in both reservoirs. Ultimately, based on the dominant genes and MAGs identified, the nitrate formed through nitrification may be reduced to back to ammonia through assimilatory process when ammonia concentrations are low. Alternatively, nitrate may be reduced to nitric oxide where it may play a role in the regulation of biofilm formation. Therefore, this information provides insight into the genetic network behind microbially mediated nitrogen metabolism in chloraminated drinking water. Here, in this project, understanding the factors that lead to potential nitrification, allowed for an in depth investigation into the metabolism of nitrogen in chloraminated drinking water, that to our knowledge has not previously been performed.

Achieving a complete understanding of the factors driving the changes in large-scale DWDS bacterial communities may be difficult, as drinking water systems are inherently dynamic and complex. However, this project has demonstrated that using the necessary technologies, the understanding of drinking water microbial ecology is improved. Here, the differences inherent to individual drinking water systems are highlighted, based on multiple contributing chemical, physical and biological factors. Understanding the microbial ecology and the factors that shape the drinking water microbiome is essential when appropriate measures to manage the microbial quality of drinking water in such a system, are to be developed, implemented and improved.

The information generated in this project highlighted the interplay between the spatial and temporal dynamics of a large-scale DWDS with multiple disinfectant regimes and demonstrated the effects of

treatment and distribution on the microbial community. This project also highlighted the importance of systematic long-term investigations to determine the seasonal effects of changes in source water quality, various environmental conditions and process operations on the microbial community dynamics. The microbial community dynamics are linked to seasonal and temporal factors such as temperature and water demand. Therefore, it is recommended that long-term, high frequency studies be performed to comprehensively understand the seasonal and annual microbial community dynamics within individual drinking water systems.

In addition, it is recommended to include absolute abundance measurements and viability assays in microbial ecology studies dealing with relative abundance data. A limitation of this study was the absence of such data and therefore it is unclear what proportion of the community were dead cells and how much of sequence data originated from extracellular DNA. Here, the lack of quantitative (i.e. bacteria concentrations) and viability data for an ecological study in an engineered system, where concentrations would likely change dramatically through disinfection and distribution, is not ideal. Including this type of data can potentially answer questions such as what causes changes in relative abundance {i.e. growth, death, DNA damage) and how many of the sequences related to dead/inactivated bacteria?

However, these investigations contribute to the current knowledge base in this field and provide the opportunity for drinking water utilities to understand the range of mechanisms that influence the microbial community structure and composition and better understand the underlying contributing factors that may cause potential challenges (e.g. nitrification). Ultimately, this growing body of information will lead to a comprehensive understanding of the drinking water microbiome that will allow the development of a drinking water environment that safeguards water quality and encourages healthy microbial communities. This will aid in creating a potential predictive framework that will facilitate proactive management of operational practices, help eliminate microbial risks and improve water quality monitoring methods.