Microbial Community Dynamics and Assembly: Drinking Water Treatment and Distribution

A Dissertation

Presented for the Doctor of Philosophy Degree

The University of Tennessee, Knoxville

Yan Zhang
December 2012

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Qiang He for his guidance through this project. I would like to thank my committee members Dr. Chris D. Cox, Dr. Gary S. Sayler and Dr. Kevin G. Robinson for their invaluable advice and assistance. I would also like to thank Dr. Alice C. Layton and Dan Williams at Center for Environmental Biotechnology for their help with pyrosequencing.

I would like to thank my coworkers in Civil and Environmental Engineering Department, Dr. Zhenwei Zhu, Si Chen and Joshua Thomas Frerichs for their help with water pipe construction and field sampling. Thank Sharon Hale for the instrument assistance. Last but not least, I want to thank my husband Hua Zeng for his support, encouragement, and unwavering belief in me. I wouldn't have made through without him.

ABSTRACT

The supply of safe drinking water is a necessity for our societies. The microbiological quality of drinking water is the most important measure of water quality directly related to public health risks from waterborne diseases. Despite the long history of water research, the processes controlling microbiological quality of drinking water remain elusive to both researchers and practitioners in the field of drinking water treatment and management, representing a critical but long standing knowledge gap. The microbial communities in drinking water systems may be influenced by multiple processes including the source water, treatment barriers, persistence to disinfection, as well as biofilm development and detachment throughout the distribution system. Previous efforts, however, are mostly limited to only one of these processes, leading to inconsistent results and incomplete understanding as expected. Taking advantages of high-throughput metagenomics tools, this research for the first time applied a systematic approach linking all relevant processes to the microbiological quality of drinking water. It is revealed that the core populations of the sampled drinking water microbial communities are dominated by *Alphaproteobacteria* and *Betaproteobacteria* affiliated to the families of *Methylobacteriaceae*, *Sphingomonadaceae*, *Comamonadaceae*, and *Oxalobacteraceae*. The characteristics of the source water and the disinfection step in the drinking water treatment process train are found to be the most important factors controlling the bacterial community structure in drinking water. Despite its potential in enhancing the removal of microbial contaminants, membrane filtration as an increasingly popular treatment alternative to rapid sand filtration is not shown to have impact differing from that of conventional rapid sand filtration on drinking water microbial communities. The compositions of drinking water microbial communities examined in this study were dominated by a few very abundant species followed by a long tail of rare species, which is well

represented by the Zipf-Mandelbrot model, accounting for 90% of the total variances and revealing low niche diversity in drinking water and distribution systems. Findings from this research provide much needed insight into the processes shaping the microbial communities in drinking water and the knowledge base for the development of effective strategies for the control of microbial contaminants in drinking water.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction

1.1. Drinking Water Treatment and Control of Microorganisms

The supply of clean drinking water is a major, and relatively recent, public health milestone (Cohn et al. 1999). Today, in most developed countries, high standards are set for drinking water quality and safety. It is well known that the presence of microorganisms in drinking water and the distribution systems causes problems of corrosion, water quality deterioration, and outbreak of waterborne diseases, of which public health risks associated with the outbreak of waterborne diseases is of the most concern (Craun et al. 2010, Li et al. 2010, White et al. 2011). Surveys conducted by the Center for Disease Control and Prevention (CDC) show that from the year 1971 to 2006, 780 outbreaks were associated with drinking water contamination, including 577,094 cases of illness and 93 deaths. Among all the outbreaks, 33% waterborne diseases were caused by contaminated source water, 39% by inadequate or interrupted treatment processes and 18% by distribution systems and premise plumbing deficiencies (Craun et al. 2010). Therefore, it is critical to understand the survival and growth of microorganisms the in drinking water and water distribution systems for the assessment, control, and prevention of public health risks associated with drinking water supply.

Publicly owned treatment works (POTWs) are the most important part of the public water supply, serving a large majority of the population in the US. Typically, drinking water treatment processes in POTWs can be separated into steps designed for the removal of suspended solids and inactivation of microorganisms. Treatment processes for solids removal typically include the sequential steps of pretreatment, coagulation, flocculation, sedimentation, and filtration, while the inactivation of microorganisms in drinking water is typically achieved through disinfection

step, as shown in Figure 1.1. Since a large fraction of the microorganisms in water present as suspended solids or attached to suspended solids, solids removal processes in drinking water treatment are as important as disinfection for the control of microbial contamination.

For surface water, pretreatment is usually a necessary step to remove large floating objects as well as certain organic, inorganic, and microbial impurities. Some water utilities pretreat the raw water using using screens, rough filters, and oxidants. These oxidizing compounds, such as permanganate and chlorine dioxide, are used to control the growth of algea, slime and other organisms, and also enhance the removal of inorganic (arsenic, iron and manganese) and natural organic compounds. Chemical coagulation is usually the next step for turbidity removal. For conventional treatment processes, coagulation is followed by flocculation and sedimentation. Slow sand filtration and some membrane filtration may be performed without coagulation. Microorganisms in natural water have similar behavior as particles and colloids as they either attached to or aggregated as particles or existing as planktonic cells. The low particle density and negative surface charges prevent microbial cells from settling out in aqueous systems. Adding coagulants will neutralize the negative charges, form inter bridges among particles and colloids, which allow them to coalesce and agglomerate into large particles, and also enmesh most of the particles and colloids in water through sweep flocculation. The most commonly used inorganic coagulants are ferric (ferric sulfate, ferric chloride) and aluminum (aluminum sulfate, aluminum chloride, aluminum hydroxide, and polyaluminum chloride) compounds. Coagulation, flocculation, and sedimentation together can typically achieve 4-log microorganism removal at optimal condition (Au 2005).

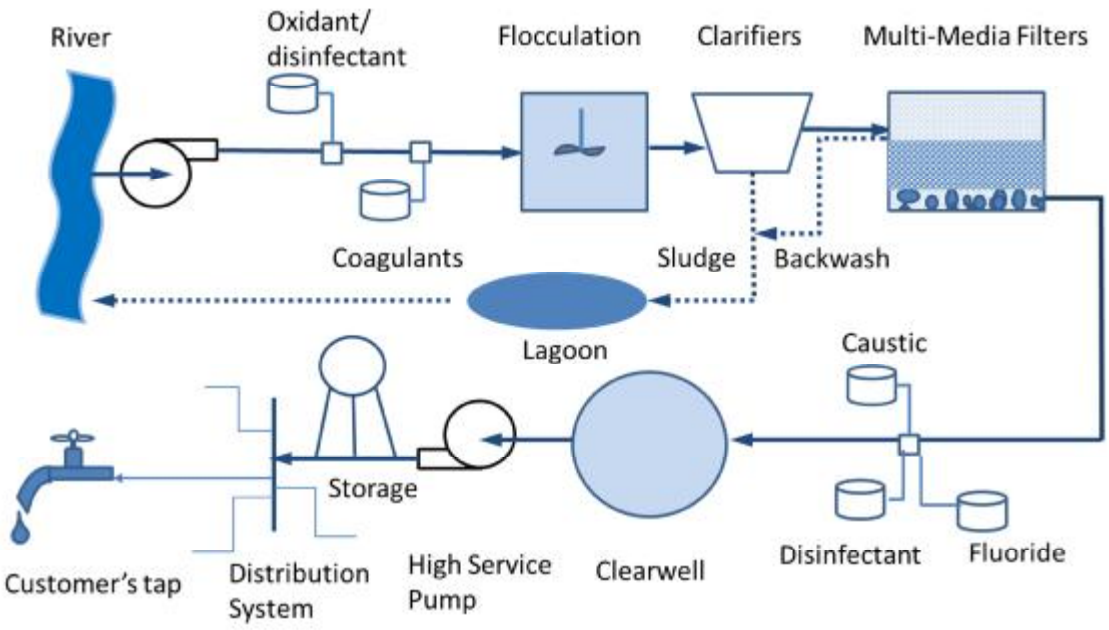

Figure 1.1 Flow diagram of conventional drinking water treatment process

Various filtration technologies are used in drinking water treatment, including slow sand filtration, rapid sand filtration, and membrane filtration. Rapid sand filtration and membrane filtration are the processes used by most Knox county utilities. The removal mechanism of membrane filtration is mainly based on the physical size of the microorganisms and the pore size of filter membrane (Figure 1.2) (AWWA and ASCE 1998). The physical size of protozoan cysts, bacteria and virus are in the range of 1 ~ 15 µm, 0.5 ~ 10 µm, and 0.02 ~ 0.08 µm, respectively. Microfiltration (MF) and ultrafiltration (UF) are designed to be able to reject particles larger than the membrane pore size (0.05 ~ 5 µm for MF and 0.005 ~ 0.05 µm for UF). Reports showed that MF and UF could achieve up to 7-log removal for particles and pathogens (Jacangelo and Watson 2002). In addition to size exclusion, transportation (diffusion, interception and sedimentation) and attachment processes are more important in rapid sand filtration. For better removal efficiency, chemical coagulation is always necessary before filtration. Results showed

that coagulation, flocculation, sedimentation and filtration together can achieve 4-log microbial removal at optimal condition (Au 2005).

Disinfection is the last treatment step for the inactivation of microorganisms in water treatment processes. Chemical oxidation is the most commonly used mechanisms in disinfection. Oxidizing agents used in the disinfection of water include chlorine and chlorine compounds, hydrogen peroxide, ozone and etc. Chlorine disinfection is performed by either applying compressed chlorine gas or sodium hypochlorite solution or calcium hypochlorite solids. When chlorine gas (Cl_2) is added into water, hypochlorous acid (HOCl) and hydrochloric acid (HCl) will form. HOCl will then partially dissociate into hypochlorite ion (OCl^-) and hydrogen ion (H^+). Chlorine also reacts with other inorganic and organic compounds in the water. The chlorine oxidation of humic and fulvic acids existing in natural water will produce trihalomethanes (THMs) and other chlorinated halides (TOX), many of which are known to be carcinogenic. The free chlorine refers to the sum of Cl_2, OCl^- and HOCl species. Chloramines can be produced by ammonia present in chlorinated water. The formation of monochloramine, dichloramine, and trichloramine depends on the chlorine/ammonia ratio, temperature, and pH. Chloramines are more stable than chlorine and result in less disinfection by-products, therefore are used by more and more utilities. The total chlorine refers to total oxidizing agents, which is the sum of free chlorine compounds and reactive chloramines. Chlorine dioxide is another alternative disinfectant used in practice to generate less disinfection by-products. Most Knoxville utilities use chlorine in primary disinfection and some utilities use sodium permanganate and chlorine dioxide in raw water pretreatment. The inactivation of microorganisms by chlorine compounds primarily involves the reaction of free chlorine with functional proteins in cell membrane and genetic materials (LeChevallier and Au 2004). However chlorine disinfectants were reported to

be ineffective for the inactivation of *Cryptosporidium* (LeChevallier and Au 2004). This concern is addressed by the increased adoption of membrane filtration technology by utilities due to the superior capability of membrane filtration in removing *Cryptosporidium* oocysts from water.

The finished water produced by water utilities is regulated to meet stringent drinking water quality standards. However, great changes may occur to finished water during its delivery from treatment plant to customers' taps due to the physical, chemical, and biological reactions in distribution system, which subsequntly lead to a potential threat to public health. Indeed, data on the outbreaks of waterborne diseases suggest that drinking water distribution systems is a source of contamination that hasn't been well studied, pointing to the urgent need to develop effective strategies to maintain water quality during the water distribution process.

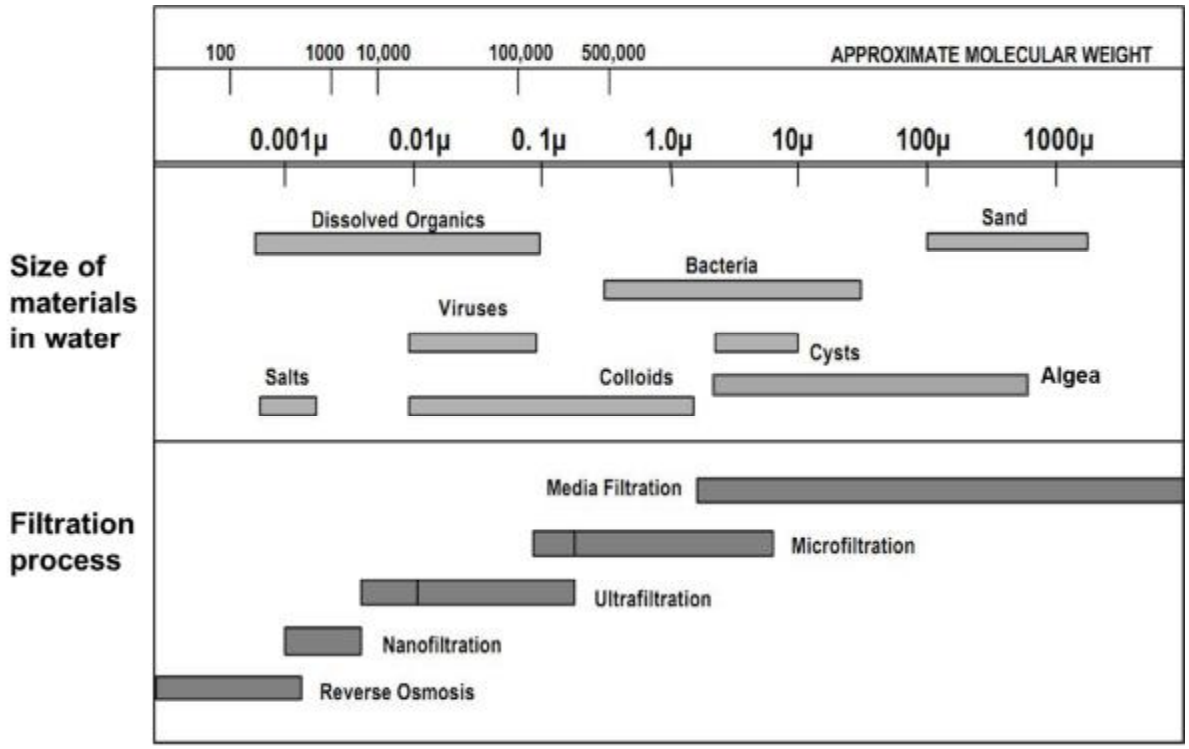

Figure 1.2 Filtration process and relative size of materials removed from water (AWWA and ASCE 1998)

The estimated total length of water distribution systems in the United States is over seven million miles, which constitute the primary management challenge for microorganism control and public health (National Research Council of the National Academies 2006). Various strategies have been explored for the control of microbial contamination in distribution systems to reduce public health risks; however, eliminating biofilm as the most prevalent and persistent form of microbial contamination remains a great challenge (Berry et al. 2006). In fact, biofilm is considered a common presence of most water distribution systems despite various controlling efforts (USEPA 2002). A major concern is that biofilms have been found to harbor and protect pathogens (Langmark et al. 2005, Saby et al. 2005, September et al. 2007, Steed and Falkinham 2006). Indeed, many different pathogens have demonstrated the ability to survive and grow, particularly in the form of biofilm, with the presence of disinfectants which is the most common approach to control microbial growth in water distribution systems (Gagnon et al. 2005, Saby et al. 2005, Torvinen et al. 2007). Thus, controlling the development of biofilm in water distribution system is a crucial step to enhance the safety of our drinking water supplies. Many factors affect the microbial growth in distribution pipelines such as temperature, disinfectant, pipe material and etc (LeChevallier et al. 1990, Pepper et al. 2004). The detachment of biofilm also affects the bulk water quality. Thus, it is important to evaluate both the treatment system and the distribution system to identify the processes contributing to the microorganisms present in the customer's tap water and associated health risks. Unfortunately, few studies have attempted to systematically identify factors impacting the microbiological quality of drinking water from source water to the end of the distribution system, representing a major knowledge gap critical for the effective assessment, control, and prevention of health risks associated with drinking water.

1.2. Factors Affect Bacteria Community in Drinking Water

1.2.1. Source water

Source water is considered to be the original source of drinking water microbial community since many of the bacteria found in tap water were found to be of fresh water origin (Henne et al. 2012, Poitelon et al. 2009). Total bacteria counts are 10^4 cell/ml in ground water and 10^7 cell/ml in surface water (Brazos and O'Connor 1984). Bacteria counts in drinking water are from 10^3 ~ 10^5 cell/ml (Hammes et al. 2008, Lautenschlager et al. 2010). USEPA standards for drinking water turbidity limit are 0.5 NTU and total bacteria no more than 500 HPC/ml in drinking water. Since the outcome of drinking water treatment is the removal and inactivation of microorganisms in source water, it is reasonable to expect that the microbes in drinking water could be traced back to the source water. Many factors including both natural and human activities influence the source water quality. Natural factors such as climate change, geology, soil run off and wild animal affect watershed. Rainfall events resulted in high levels of turbidity in Delaware River was reported to lead to peak levels of microbial contaminants (LeChevallier et al. 1998). Seasonal increases of coliform bacteria in some northeastern watersheds in the USA was considered to be caused by seasonal migratory of birds (Robbins et al. 1991). Human activities including wastewater discharges, livestock, and recreational activities, were considered major sources of surface water contamination. Municipal wastewater and livestock were considered as major sources for *Bacteroides* species and coliform bacteria in watersheds in Tennessee (Layton et al. 2006). Recreational activities involving body contact was also a source for fecal contamination (Stewart et al. 2002).

1.2.2. Filtration technology

The application of microfiltration (MF) and ultrafiltration (UF) has been rapidly increasing in recent years due to their superior capacity for the removal of particles and microorganisms (Alspach et al. 2008, Jacangelo and Watson 2002). Many reports show that MF and UF could achieve up to 7-log removal for particles and pathogens (Jacangelo and Watson 2002). Both lab scale and pilot scale studies have shown that MF could act as an absolute barrier to protozoan cysts (LeChevallier and Au 2004). Through MF and UF *Giardia muris*, *Cryptosporidium parvum*, total coliforms, *Escherichia coli* and enterococci in the filtered water were all below detection limit (Jacangelo et al. 1991, Jacangelo et al. 1995). Some literature showed that conventional filtration combined with coagulation, flocculation, sedimentation together can achieve 4-log microbial removal at optimal condition (Au 2005). Lab scale comparison suggested that membrane filtration had better efficiency for microorganism removal than conventional filtration process and generate a different microbial community structure (Ho et al. 2012). In theory MF and NF could exclude any microorganism larger than the pore size including protozoa, algae and most bacteria. However HPC was still detected in some cases (Jacangelo et al. 1991). The defects in membrane fiber may allow microbes to escape membrane barrier. The attached microorganisms may cause fouling problem (Guo et al. 2010). However most previous studies only targeted the removal performance of total cell numbers and several specific pathogens, the whole microbial community changes after filtration haven't been thoroughly investigated. More importantly, nothing is known about the influence of membrane filtration on drinking water microbial community as compared to conventional filtration.

1.2.3. Disinfection

There are several factors that affect disinfection efficiency: disinfectant type and concentration, contact time, temperature, pH, and the presence of organic matter and particles. The CT concept (disinfectant concentration times the contact time) is used to evaluate disinfection efficiency. USEPA also set up regulation rules for different pathogens based on the CT value (USEPA 1991). Previous studies showed that disinfection efficiency order for disinfectant were ozone > chlorine dioxide > free chlorine > chloramines. However their stability in water was shown as opposite order ozone < free chlorine < chlorine dioxide < chloramines (Aieta and Berg 1986).

Ehrlick et al. suggested that a requirement for chlorine residuals to continued suppression of microbial growth in underground water is around 2.5 mg/l (Ehrlich et al. 1979). The water pH is another important factor for bacteria inactivation. Free chlorine had better disinfection efficiency at lower pH (6.0 ~7.0) than at high pH (8.5 ~ 10.0). Monochloramine formation occurs in the pH range of 7~ 9 and CT value for heterotrophic bacteria inactivation is lower at pH 7.0 than pH 8.5. While chlorine dioxide had better performance at alkaline pH levels (7.0 ~ 8.5) (Lechevallier et al. 1988). Merely increasing disinfectant concentration may not increase disinfection efficiency because other factors such as organic matter, particles and biofilm also affect disinfection efficiency as they are both food resources for microorganisms and causes for disinfectant decay. Kooij suggested that HPC may be limited when assimilable organic carbon (AOC) < 50 ug/L (Van der Kooij and Hijnen 1985). Lechevallier suggested that TOC content threshold for coliform bacteria occurrence is 2.4mg/L (LeChevallier et al. 1991). The attachment to particles production of the extracellular capsule and aggregation of bacteria in drinking water can also affect disinfection effciency. The aggregation of *Acinetobacter* showed 2.3 fold more resistance to monochloramine and 100 fold more resistance to hypochlorous acid

(Stewart and Olson 1986). The application of chloramine promoted the growth of nitrifiers in drinking water (Eichler et al. 2006). Disinfection was reported to be the key step shaping the bacteria community structure in drinking water with *Proteobacteria* present as the dominate population (Eichler et al. 2006, Poitelon et al. 2010). Some results showed that chlorination caused bacterial population to shift from gram-negative to gram-positive (Norton and LeChevallier 2000, Pepper et al. 2004).

1.2.4. Distribution system

No matter how stringent the drinking water is treated, some microorganisms can always pass the treatment barriers and persist in the drinking water distribution systems (National Research Council of the National Academies 2006). Many microorganisms can attach to the interior surface of pipelines and formed biofilm in drinking water distribution systems (USEPA 2002). Biofilms are suspected to be a major source of microorganisms in distribution systems that carry adequately treated water with no pipeline defects (Berry et al. 2006, Lechevallier et al. 1987). Indeed, recent studies have found that the majority (~ 95%) of the overall biomass in distribution systems is present in biofilms attached to the pipe surfaces (Flemming et al. 2002, Servais et al. 2004). Biofilms predominate because cells in the biofilm matrix may have certain advantages over planktonic cells, such as increased protection from disinfection (National Research Council of the National Academies 2006). This is supported by findings that the presence of disinfectants is effective in reducing the concentration of planktonic bacteria, but has little or no effect on the concentration of biofilm bacteria (Gagnon et al. 2005, Kuo and Chen 2006). The mechanism behind the observed prevalence and persistence of biofilm in the harsh environment of distribution systems is still unknown (Berry et al. 2006), although hypotheses include disinfectants mass transfer resistance (Chambless et al. 2006, Stewart et al. 1996), the formation

of persister cells (Lewis 2005, Roberts and Stewart 2005), and protection from the produced extracellular polymeric substances (Flemming et al. 2007, Samrakandi et al. 1997). However, these hypotheses may be fundamentally flawed as they are mostly developed from pure culture models, instead of multispecies microbial communities, where interspecies interactions are more important for survival and adaptation (Simoes et al. 2007). Therefore, the community level approach is required to identify the microbial processes underlying the prevalence and persistence of biofilms in distribution systems so that effective biofilm control strategies targeting these critical processes could be developed.

Literatures showed that many factors affected microbial growth in drinking water distribution systems including pipe material, disinfectant type and concentration, physical and chemical characteristics of bulk water and etc. LeChevallier and Mathieu et al studied the influence of disinfectant type on the growth of biofilm (LeChevallier et al. 1990, Mathieu et al. 1993). However, their results were not consistent to each other. LeChevallier's results showed that chloramine had a stronger ability to penetrate biofilm than free chlorine. While Mathieu's experiment drew an opposite conclusion showing that chlorine was more effective than chloramine. Many researchers studied the effect of pipe material on bacteria growth and microbial community composition. Henne et al studied biofilm on steel, copper, plastic, glass, and Teflon surface from drinking water systems through single strand confirmation polymorphism (SSCP). They found that *Alphaproteobacteria* (26%) was the most dominant population (Henne et al. 2012). Jang et al also found that *Alphaproteobacteria* dominant on both copper and PVC pipe surface based on DGGE analysis (Jang et al. 2011). Pavissich' results showed that *Betaproteobacteria* and *Gammaproteobacteria* were the dominant population on copper coupons based on T-RFLP analysis (Pavissich et al. 2010). Schwartz found

Betaproteobacteria and *Gammaproteobacteria* were higher than *Alphaproteobacteria* on PVC, hardened PE and steel surface (Schwartz et al. 2003). Kalmbach studied the effect of glass, low and high density polyethylene, and soft PVC on bacterial community and found that *Betaproteobacteria* was the dominant population. The bacterial community composition on soft-PVC was different significantly from other materials, which was dominated by *Aquabacterium citratiphilum* while *Aquabacterium commune* dominated on other pipe surface (Kalmbach et al. 2000). However, the results from different studies were not consistant. And most of the previous studies have not provided sufficient information to tell the best pipe material for plumbing system. Therefore, in this study we chose three commonly used pipe material in Knox County and investigate the influence of pipe material on the biofilm microbial community composition though pyrosequencing analysis.

Since most water treatment plants in local area use chlorine as disinfectant, in this study we focus on the influence of three different pipe materials on bacteria biofilm attached in the pipe wall. The water main pipe connected to water treatment plant is usually 24-in to 8-inch cast iron. The service pipelines close to customers' end are usually made of three different pipe materials in Knox County: galvanized iron, copper and polyvinyl chloride (PVC). Water pipes in old buildings and houses were mainly galvanized iron pipes. Buildings for commercial usage chose copper pipes for their drinking water supply systems. New houses built in recent year tend to use PVC pipe instead of copper pipes to cut the expenses. Therefore in our study we set up three different pipe loops using the following pipe materials: galvanized iron, copper, and polyvinyl chloride (PVC) to mimic premise plumbing networks.

1.3. Methods for Studying Microbial Community in Drinking Water

The concept of microbial ecology in drinking water distribution system was first articulated by Wilson (Wilson 1945). He suggested that bacterial type and number developed in drinking water distribution system depend on the available ecological niche. The understanding of microbial community in drinking water started from the bacteria cultures and pathogens isolated from water. Heterotrophic bacterial count, coliform count and some pathogen detection were all culture base methods. However majority of microorganisms cannot be cultured in a laboratory. With the development of molecular techniques the difficulty for cultivation was circumvented by DNA extraction from environmental sample and the understanding of microbial ecology went deep into molecular level. Real time PCR probes were developed for trace pathogen detection (Wang et al. 2012). Many finger print molecular techniques such as clone library, denaturing gradient gel electrophoresis (DGGE), single strand conformation polymorphism (SSCP), fluorescence in situ hybridization (FISH), and terminal restriction fragment length polymorphism (T-RFLP) were used to study the microbial community composition in drinking water. Revetta and Poitelon studied the drinking water bacterial community using 16S rRAN and rRNA gene based clone library. They found that the most dominant population was difficult to be classified suggesting that there might be novel bacteria lineages in drinking water that haven't been observed (Poitelon et al. 2009, Revetta et al. 2010). Hoefel used DGGE and FISH to reveal the occurrence of nitrifying bacteria that were associated with the appliance of chloramine (Hoefel et al. 2005). Ho's result based on DGGE showed that membrane filtration generated different bacterial communities than other conventional treatment systems (Ho et al. 2012). Eichler utilized SSCP technique to show the contribution of source water and chloramines to the drinking water bacterial community by showing the composition of fresh water origin bacteria

and nitrifying bacteria (Eichler et al. 2006). Henne investigated bulk water and biofilm in drinking water distribution system using DGGE fingerprint and revealed different bacterial communities for bulk water and biofilm (Henne et al. 2012). He also concluded that the physical location had more influence on bacterial community in biofilm than pipe material did. However the previous studies were based on case studies. To understand the microbial ecology in drinking water, a systematic study needs to be done.

The application of next generation sequencing in recent years greatly increased the sampling depth. Hong's investigation of water meter biofilm revealed a broad variety of bacteria that may live on methane and other naturally occurring compounds. There are a number of different platforms for massively parallel DNA sequencing: Roche/454 FLX, Illumina/Solexa Genome Analyzer, Applied Biosystems SOLiD™ System, Helicos Heliscope™, and Pacific Biosciences SMRT. 454 pyrosequencing was the most popular techniques in the past several years due to the advantages in the relative long sequence and high throughput. Pinto's study showed that drinking water microbial community was governed by filtration process (Pinto et al. 2012). For our study we use 454 pyrosequencing for bacterial community analysis since this method had better coverage and resolution.

There are advantages as well as disadvantages for pyrosequencing. Compared with Sanger sequencing one of the greatest advantages of 454 pyrosequencing is that hundreds of thousands of sequence reads can be generated in a single run, thus the cost per base is much lower than Sanger sequencing. In addition, sequences from different samples can be labeled through multiplex barcode approach and pooled together for the same run, which subsequently increases efficiency and reduces costs. 454 the pyrosequencing method greatly increased sampling depth, coverage and the chances for detecting rare species. The sampling depth for

clone library depends on the number of clones that were picked and sequenced. One run for 454 FLX titanium gives 700k reads (454.com). Another advantage of the pyrosequencing technique is that the tedious bacterial cloning steps are skipped and therefore the clone related biases are avoided.

One disadvantage of pyrosequencing approach is the detection of long homopolymers, which results in sequencing errors. The sequencing errors may be recognized as new rare operational taxonomic units (OTU), which consequently results in misleading the diversity and richness estimates. Several denoising approaches based on both signal flow and sequence analysis have been proposed to deal with sequencing errors (Huse et al. 2010, Quince et al. 2009). The sequencing length generated by 454 pyrosequencing (250 ~ 400bp) is shorter than Sanger sequencing which give 700 ~ 900bp per reaction. The short reads limits the bacteria classification into deeper phylogenetic level.

1.4. Community Assembly Theory and Relative Abundance Distribution

One of the central goals in community ecology is to understand the processes that structure biological communities. Many environmental factors, as well as biological interactions such as species adaptation, competition, evolution, and dispersion, could be involved in the microbial community assembly process. Subsequently, microbial communities may display different structure patterns in both natural and engineered environments (Inceoglu et al. 2011, Schloss and Handelsman 2006, Sloan et al. 2007). Relative abundance distribution (RAD) and species abundance distribution (SAD) are the two major ways to illustrate community structure (McGill et al. 2007). The RAD has been widely used to compare the structure differences since the SAD curves have been considered as biased by the arbitrary abundance categories (Wilson 1991). Except for the observation and description of RAD, all researches expected more ecological

information and sought an answer to the question about how this specific community structure was shaped. Ecologists generalized all kinds of RAD patterns and developed more than 40 ecological models to integrate RAD with ecological processes. The rationale behind this effort is that RAD patterns are formed from specific community assembly processes; therefore, the RADs might indicate the corresponding process. Thus, the mathematical model derived from ecological process might be able to predict a particular community structure. One possible way to link the observed RAD to its ecological context is to find the best fit ecological model for the distribution curve and explain the community assembly process through the best fit model.

Many stochastic and deterministic hypotheses have been proposed to describe relative abundance distribution and explain the assembling mechanisms for organisms in natural environments (Hubbell 2001, Tokeshi 1993). Basically four categories of model are used to describe species abundance distribution pattern: 1) statistic model (eg. lognormal and logseries); 2) niche partitioning model (eg. broken stick, geometric series and Tokeshi's models); 3) population dynamics (eg. neutral model); 4) branching process (eg. Zipf-Mandelbrot and fractal branching models). These models have been used extensively to address the ecological rules that govern the diversity and abundance of plant and animal communities (Hubbell 2001, Watkins and Wilson 1994). Most ecological models were extensively tested and used in macro ecology to address the ecological rules that may govern the assembly of plant and animal communities (Hubbell 2001, Watkins and Wilson 1994). With the development of molecular techniques, the whole community, instead of only the culturable microbes, can be sampled and investigated. Till very recent years, microbial ecologists managed to extrapolate some models from macro ecological to micro ecology (Horner-Devine et al. 2004, Prosser et al. 2007). The development of next generation sequencing has significantly improved the coverage of bacterial community in

environmental samples, particularly rare populations that could not be readily identified by other techniques. The massive information derived from these high-throughput methods has allowed us to test the macro ecological theory in micro scale. A few studies focused on the bacterial community structures in soil, ocean and even wastewater (Galand et al. 2009, Schloss and Handelsman 2006, Sloan et al. 2007). However, no reports on the application of these models on drinking water communities are available in the literature. Thus, modeling efforts are needed to potentially identify factors controlling the bacteria community composition in the drinking water.

The application of pyrosequencing technology in drinking water bacteria provides a comprehensive view of microbial assemblages, revealing the presence of both abundant and rare species (Hong et al. 2010, Kwon et al. 2011). Its deep sampling advantage also provides a great opportunity to explore the ecological process underlying microbial diversity patterns by characterizing the whole species abundance distribution pattern, instead of merely focusing on the predominant species. Thus in this study we will use high-throughput sequencing technology to gain more comprehensive understanding of the microbial communities in drinking water.

1.5. Statement of Problems and Study objectives

1.5.1. Problems

It is of great importance to identify the factors controlling drinking water microbial quality, which requires an understanding of the entire microbial community in drinking water. However, most previous studies relied on cultivation-based techniques, such as heterotrophic plate counts (HPC) and the counting of specific pathogens. The majority of microorganisms in drinking water are likely not culturable, these studies are apparently biased. To overcome this bias, recent efforts have used, cultivation-independent techniques, primarily molecular

techniques such as clone library and fingerprint analysis, for the characterization of microbial populations in drinking water, revealing remarkable bacterial diversity and dynamics in drinking water through the analysis of 16S rRNA gene sequences. However, fingerprints based detection methods were considered limited by their sampling depth, since only abundant populations were sampled by these methods. The development of next-generation DNA sequencing technologies, capable of unprecedented sampling capacity, provides a potential opportunity to explore the large bacterial diversity in drinking water. Therefore, it is important to use high-throughput pyrosequencing analysis to gain a full understanding of drinking water microbial community.

Another challenge to understand the microbial ecology of drinking water is the lack of systematic investigation of the entire drinking water process train, since the microbial community structure is influenced by multiple steps during treatment and the distribution system. Because almost all previous studies were focused on water samples taken from individual treatment steps or single points in the distribution system great variances in microbial community composition were observed in different studies. The results from different studies were not comparable since the source water, treatment techniques, and pipe materials in the distribution system were all different. Systematic investigation of the whole drinking water treatment and distribution system was necessary to identify the treatment steps with the most significant influences on the microbial ecology of drinking water. Therefore this is a major knowledge gap requiring the understanding of changes in the microbial community along the treatment process train from source water to the customer's tap.

Filtration is a critical treatment step in the control of microbiological quality of drinking water. Membrane filtration has been adopted as the central alternative to conventional rapid sand filtration to meet increasingly stringent water quality standards, particularly for the removal of

Cryptosporidium. However, it remains virtually unknown how membrane filtration influences the microbiological quality of water in the distribution system and at the tap. Therefore, an accurate assessment of the influence of membrane filtration requires the understanding of the changes in microbial ecology throughout the entire process train. In this study we selected four water treatment plants including two membrane filtration plants and two conventional filtration plants taking two different rivers as source water. So that for each source water comparisons can be made between one conventional filtration plant and one membrane filtration plant.

Previous studies of the microbiological quality of drinking water have provided extensive but inconsistent information on the patterns of the microbial community dynamics in drinking water. The discrepancies in the patterns of community structure discussed above suggest that our understanding of the processes driving community assembly remains incomplete, which could be attributed to the lack of systematic profiling of the microbial community throughout the entire treatment process train and the distribution system. Built upon the systematic characterization of microbial communities in multiple drinking water treatment utilities from the source water to the tap in this research, which provided near-complete coverage of the microbial community compositions required for the modeling of community dynamics using high-throughput sequencing, we attempt to quantitatively characterize the microbial community distribution pattern using ecological models. This effort will provide much needed insight into the major processes shaping the microbial community structure in drinking water, which is critical for the development of effective strategies for the control of microbial contamination in drinking water and minimize public health risks.

1.5.2. Objectives

Drinking water treatment is one of the most important environmental engineering processes in operation worldwide. The microbial assemblages in drinking water distribution systems are believed to be a potential reservoir for bacteria which may be a threat for the water quality and subsequently human health. The applications of molecular techniques in drinking water microbial community revealed a broad diversity of microorganisms. The factors that may affect drinking water microbial community were discussed through the above literature review. Based on previous studies we developed the following hypotheses: 1) drinking water host diverse bacteria seeded by the source river and biofilms drinking water distribution system; 2) among the following factors: source water, treatment technology, disinfection, and distribution, disinfection play the most important role in shaping drinking water bacterial community; 3) compared to conventional filtration, membrane filtration has better treatment efficiency and produces different bacterial communities; 4) biofilm bacterial communities are affected by pipe materials.

To investigate the above hypotheses and to understand the determinants of microbial community structures in drinking water treatment and distribution systems, the primary objectives of this study are to 1) gain a systematic understanding of the dynamics of microbial communities in drinking water treatment and distribution processes; 2) identify the major factors controlling the microbial community structure in drinking water; 3) evaluate membrane filtration as a critical technological alternative on the microbiological quality of drinking water; 4) characterize the patterns of microbial community assembly in drinking water.

References

Aieta, E.M. and Berg, J.D. (1986) A Review of Chlorine Dioxide in Drinking Water Treatment. Journal American Water Works Association 78(6), 62-72.

Alspach, B., Adham, S., Cooke, T., Delphos, P., Garcia-Aleman, J., Jacangelo, J., Karimi, A., Pressman, J., Schaefer, J., Sethi, S. and Publicatic, A.S.P. (2008) Microfiltration and ultrafiltration membranes for drinking water. Journal American Water Works Association 100(12), 84-97.

Au, K.-K. (2005) Water Encyclopedia, John Wiley & Sons, Inc.

AWWA and ASCE (1998) Water Treatment Plant Design. Third Edition, AWWA and ASCE, McGraw-Hill, Inc, New York, NY.

Berry, D., Xi, C.W. and Raskin, L. (2006) Microbial ecology of drinking water distribution systems. Current Opinion in Biotechnology 17(3), 297-302.

Brazos, B.J. and O'Connor, J.T. (1984) ENUMERATION METHODS FOR BACTERIA IN WATER DISTRIBUTION-SYSTEMS. Journal American Water Works Association 76(4), 60-60.

Chambless, J.D., Hunt, S.M. and Stewart, P.S. (2006) A three-dimensional computer model of four hypothetical mechanisms protecting biofilms from antimicrobials. Applied and Environmental Microbiology 72(3), 2005-2013.

Cohn, P.D., Cox, M. and Berger, P.S. (1999) Health and aesthetic aspects of water quality. In American Water Works Association (ed.), Water quality and treatment: a handbook of community water supplies, 5th ed. , McGraw Hill, Inc., New York, N. Y.

Craun, G.F., Brunkard, J.M., Yoder, J.S., Roberts, V.A., Carpenter, J., Wade, T., Calderon, R.L., Roberts, J.M., Beach, M.J. and Roy, S.L. (2010) Causes of Outbreaks Associated with Drinking Water in the United States from 1971 to 2006. Clinical Microbiology Reviews 23(3), 507-528.

Ehrlich, G.G., Ku, H.F.H., Vecchioli, J. and A., E.T. (1979) Microbiological effects of recharging the Magothy aquifer, Bay Park, New York, with tertiary-treated sewage: deep-well artifical recharge experiments at Bay Park Long Island, New York. United States Geological Survey. Professional paper 751-E.

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.

Flemming, H., Percival, S. and Walker, J. (2002) Contamination potential of biofilms in water distribution systems. Water Supply 2(1), 271-280.

Flemming, H.C., Neu, T.R. and Wozniak, D.J. (2007) The EPS matrix: The "House of Biofilm cells". Journal of Bacteriology 189(22), 7945-7947.

Gagnon, G.A., Rand, J.L., O'Leary, K.C., Rygel, A.C., Chauret, C. and Andrews, R.C. (2005) Disinfectant efficacy of chlorite and chlorine dioxide in drinking water biofilms. Water Research 39(9), 1809-1817.

Galand, P.E., Casamayor, E.O., Kirchman, D.L. and Lovejoy, C. (2009) Ecology of the rare microbial biosphere of the Arctic Ocean. Proceedings of the National Academy of Sciences of the United States of America 106(52), 22427-22432.

Guo, H., Wyart, Y., Perot, J., Nauleau, F. and Moulin, P. (2010) Low-pressure membrane integrity tests for drinking water treatment: A review. Water Research 44(1), 41-57.

Hammes, F., Berney, M., Wang, Y.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 42(1-2), 269-277.

Henne, K., Kahlisch, L., Brettar, I. and Hofle, 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.

Ho, L., Braun, K., Fabris, R., Hoefel, D., Morran, J., Monis, P. and Drikas, M. (2012) Comparison of drinking water treatment process streams for optimal bacteriological water quality. Water Research 46(12), 3934-3942.

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.

Hong, P.Y., Hwang, C.C., Ling, F.Q., Andersen, G.L., LeChevallier, M.W. and Liu, W.T. (2010) Pyrosequencing Analysis of Bacterial Biofilm Communities in Water Meters of a Drinking Water Distribution System. Applied and Environmental Microbiology 76(16), 5631-5635.

Horner-Devine, M.C., Lage, M., Hughes, J.B. and Bohannan, B.J.M. (2004) A taxa-area relationship for bacteria. Nature 432(7018), 750-753.

Hubbell, S.P. (2001) The unified neutral theory of biodiversity and biogeography, Princeton University Press, Princeton.

Huse, S.M., Welch, D.M., Morrison, H.G. and Sogin, M.L. (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environmental Microbiology 12(7), 1889-1898.

Inceoglu, O., Abu Al-Soud, W., Salles, J.F., Semenov, A.V. and van Elsas, J.D. (2011) Comparative Analysis of Bacterial Communities in a Potato Field as Determined by Pyrosequencing. Plos One 6(8).

Jacangelo, J.G., Laine, J.M., Carns, K.E., Cummings, E.W. and Mallevialle, J. (1991) Low-pressure membrane filtration for removing Giardia and microbial indicators. American Water Works Association Journal 83(9), 97-106.

Jacangelo, J.G., Adham, S.S. and Laine, J.-M. (1995) Mechanism of Cryptosporidium, Giardia, and MS2 virus removal by MF and UF. American Water Works Association Journal 87(9), 107-121.

Jacangelo, J.G. and Watson, M. (2002) Control of microorganisms in drinking water, American Society of Civil Engineers, Reston, Virginia.

Jang, H.J., Choi, Y.J. and Ka, J.O. (2011) Effects of Diverse Water Pipe Materials on Bacterial Communities and Water Quality in the Annular Reactor. Journal of Microbiology and Biotechnology 21(2), 115-123.

Kalmbach, S., Manz, W., Bendinger, B. and Szewzyk, U. (2000) In situ probing reveals Aquabacterium commune as a widespread and highly abundant bacterial species in drinking water biofilms. Water Research 34(2), 575-581.

Kuo, J.F. and Chen, J. (2006) Disinfection antimicrobial and processes. Water Environment Research 78(10), 1406-1428.

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.

Langmark, J., Storey, M.V., Ashbolt, N.J. and Stenstrom, T.A. (2005) Biofilms in an urban water distribution system: measurement of biofilm biomass, pathogens and pathogen persistence within the Greater Stockholm area, Sweden. Water Science and Technology 52(8), 181-189.

Lautenschlager, K., Boon, N., Wang, Y.Y., Egli, T. and Hammes, F. (2010) Overnight stagnation of drinking water in household taps induces microbial growth and changes in community composition. Water Research 44(17), 4868-4877.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R. and Sayler, G. (2006) Development of Bacteroides 16S rRNA Gene TaqMan-Based Real-Time PCR Assays for Estimation of Total, Human, and Bovine Fecal Pollution in Water. Applied and Environmental Microbiology 72(6), 4214-4224.

Lechevallier, M.W., Babcock, T.M. and Lee, R.G. (1987) Examination and characterization of distribution system biofilms. Applied and Environmental Microbiology 53(12), 2714-2724.

Lechevallier, M.W., Cawthon, C.D. and Lee, R.G. (1988) Factors promoting survival of bacteria in chlorinated water supplies. Applied and Environmental Microbiology 54(3), 649-654.

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

LeChevallier, M.W., Schulz, W. and Lee, R.G. (1991) Bacterial nutrients in drinking water. Applied and Environmental Microbiology 57(3), 857-862.

LeChevallier, M.W., Norton, W.D., Abbaszadegan, M., Atherholt, T.B. and Rosen, J. (1998) Development of a Monitoring Strategy to Determine Variations in Giardia and Cryptosporidium Levels in a Watershed, Dallas, TX. .

LeChevallier, M.W. and Au, K. (2004) Water treatment and pathogen control: process efficiency in achieving safe drinking-water.

Lewis, K. (2005) Persister cells and the riddle of biofilm survival. Biochemistry-Moscow 70(2), 267-+.

Li, D., Li, Z., Yu, J., Cao, N., Liu, R. and Yang, M. (2010) Characterization of bacterial community structure in a drinking water distribution system during an occurrence of red water. Appl. Environ. Microbiol. 76(21), 7171-7180.

Mathieu, L., Block, J.C., Dutang, M., Maillard, L. and Reasoner, D. (1993) Control of biofilm accumulation in drinking-water distribution systems. Water Supply 11(3/4), 365-376.

McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., Dornelas, M., Enquist, B.J., Green, J.L., He, F.L., Hurlbert, A.H., Magurran, A.E., Marquet, P.A., Maurer, B.A., Ostling, A., Soykan, C.U., Ugland, K.I. and White, E.P. (2007) Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology Letters 10(10), 995-1015.

National Research Council of the National Academies (2006) Drinking Water Distribution Systems: Assessing and Reducing Risks., National Academies Press, Washington, D.C.

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.

Pavissich, J.P., Vargas, I.T., González, B., Pastén, P.A. and Pizarro, G.E. (2010) Culture dependent and independent analyses of bacterial communities involved in copper plumbing corrosion. Journal of Applied Microbiology 109(3), 771-782.

Pepper, I.L., Rusin, P., Quintanar, D.R., Haney, C., Josephson, K.L. and Gerba, C.P. (2004) Tracking the concentration of heterotrophic plate count bacteria from the source to the consumer's tap. International Journal of Food Microbiology 92(3), 289-295.

Pinto, A., Xi, C. and Raskin, L. (2012) Bacterial community structure in the drinking water microbiome is governed by filtration processes. Environmental Science & Technology.

Poitelon, J.B., Joyeux, M., Welte, B., Duguet, J.P., Prestel, E., Lespinet, O. and Dubow, M.S. (2009) Assessment of phylogenetic diversity of bacterial microflora in drinking water using serial analysis of ribosomal sequence tags. Water Research 43(17), 4197-4206.

Poitelon, J.B., Joyeux, M., Welte, 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 & Biotechnology 37(2), 117-128.

Prosser, J.I., Bohannan, B.J.M., Curtis, T.P., Ellis, R.J., Firestone, M.K., Freckleton, R.P., Green, J.L., Green, L.E., Killham, K., Lennon, J.J., Osborn, A.M., Solan, M., van der Gast, C.J. and Young, J.P.W. (2007) Essay - The role of ecological theory in microbial ecology. Nature Reviews Microbiology 5(5), 384-392.

Quince, C., Lanzen, A., Curtis, T.P., Davenport, R.J., Hall, N., Head, I.M., Read, L.F. and Sloan, W.T. (2009) Accurate determination of microbial diversity from 454 pyrosequencing data. Nature Methods 6(9), 639-U627.

Revetta, R.P., Pemberton, A., Lamendella, R., Iker, B. and Santo Domingo, J.W. (2010) Identification of bacterial populations in drinking water using 16S rRNA-based sequence analyses. Water Research 44(5), 1353-1360.

Robbins, R.W., Glicker, J.L., Bloem, D.M. and Niss, B.M. (1991) Effective watershed management for surface water supplies. American Water Works Association Journal 83(12), 34-44.

Roberts, M.E. and Stewart, P.S. (2005) Modelling protection from antimicrobial agents in biofilms through the formation of persister cells. Microbiology-Sgm 151, 75-80.

Saby, S., Vidal, A. and Suty, H. (2005) Resistance of Legionella to disinfection in hot water distribution systems. Water Science and Technology 52(8), 15-28.

Samrakandi, M.M., Roques, C. and Michel, G. (1997) Influence of trophic conditions on exopolysaccharide production: bacterial biofilm susceptibility to chlorine and monochloramine. Canadian Journal of Microbiology 43(8), 751-758.

Schloss, P.D. and Handelsman, J. (2006) Toward a census of bacteria in soil. Plos Computational Biology 2(7), 786-793.

Schwartz, T., Hoffmann, S. and Obst, U. (2003) Formation of natural biofilms during chlorine dioxide and u.v. disinfection in a public drinking water distribution system. Journal of Applied Microbiology 95(3), 591-601.

September, S.M., Els, F.A., Venter, S.N. and Brozel, V.S. (2007) Prevalence of bacterial pathogens in biofilms of drinking water distribution systems. Journal of Water and Health 5(2), 219-227.

Servais, P., Anzil, A., Gatel, D. and Cavard, J. (2004) Biofilms in the Parisian suburbs drinking water distribution system. J Water SRT-Aqua 53.

Simoes, L.C., Simoes, M. and Vieira, M.J. (2007) Biofilm interactions between distinct bacterial genera isolated from drinking water. Applied and Environmental Microbiology 73(19), 6192-6200.

Sloan, W.T., Woodcock, S., Lunn, M., Head, I.M. and Curtis, T.P. (2007) Modeling taxa-abundance distributions in microbial communities using environmental sequence data. Microbial Ecology 53(3), 443-455.

Steed, K.A. and Falkinham, J.O. (2006) Effect of growth in biofilms on chlorine susceptibility of Mycobacterium avium and Mycobacterium intracellulare. Applied and Environmental Microbiology 72(6), 4007-4011.

Stewart, M. and Olson, B. (1986) Mechanisms of bacterial resistance to inorganic chloramines, Portland, OR.

Stewart, M.H., Yates, M.V., Anderson, M.A., Gerba, C.P., Rose, J.B., Leon, R.D. and Wolfe, R.L. (2002) Predicted Public Health Consequences of Body-contact Recreation ON A POTABLE WATER RESERVOIR. Journal (American Water Works Association) 94(5), 84-97.

Stewart, P.S., Hamilton, M.A., Goldstein, B.R. and Schneider, B.T. (1996) Modeling biocide action against biofilms. Biotechnology and Bioengineering 49(4), 445-455.

Tokeshi, M. (1993) Species abundance patterns and community structure. Advances in Ecological Research, Vol 24 24, 111-186.

Torvinen, E., Lehtola, M.J., Martikainen, P.J. and Miettinen, I.T. (2007) Survival of Mycobacterium avium in drinking water biofilms as affected by water flow velocity, availability of phosphorus, and temperature. Applied and Environmental Microbiology 73(19), 6201-6207.

USEPA (1991) Guidance Manual for Compliance with the Filtration and Disinfection Requirements for Public Water Systems Using Surface Water Sources. Agency, U.S.E.P. (ed), Washington DC.

USEPA (2002) Health risks from microbial growth and biofilms in drinking water distribution systems, Distribution System Issue paper., Washington D.C.

Van der Kooij, D. and Hijnen, W. (1985) Measuring the concentration of easily assimilable organic carbon in water treatment as a tool for limiting regrowth of bacteria in distribution systems.

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.

Watkins, A.J. and Wilson, J.B. (1994) Plant community structure, and its relation to the vertical complexity of communities: dominance/diversity and spatial rank consistency. Oikos 70(1), 91-98.

White, C., Tancos, M. and Lytle, D.A. (2011) Microbial Community Profile of a Lead Service Line Removed from a Drinking Water Distribution System. Applied and Environmental Microbiology 77(15), 5557-5561.

Wilson, C. (1945) Bacteriology of water pipes. Jour Amer Water Works Assoc 37((1)), 52-58.

Wilson, J.B. (1991) Methods for Fitting Dominance Diversity Curves. Journal of Vegetation Science 2(1), 35-46.

Chapter 2. Characterization of Bacterial Diversity in Drinking Water by Pyrosequencing

A version of this chapter was originally published by Yan Zhang and Qiang He:
Yan Zhang and Qiang He. “Characterization of bacterial diversity in drinking waterby pyrosequencing” *Water Science & Technology: Water Supply* 2012 (in press).

2.1. Abstract

Controlling microbial contamination of drinking water is critical to public health. However, understanding of the microbial ecology of drinking water remains incomplete. Representing the first application of high-throughput sequencing in drinking water microbiology, the objective of this study is to evaluate pyrosequencing as a high-throughput technique for the characterization of bacterial diversity in drinking water in comparison with conventional clone library analysis. Pyrosequencing and clone library analysis were performed in parallel to study the bacterial community composition in drinking water samples following the concentration of microbial biomass in drinking water with ultrafiltration. Validated by clone library analysis, pyrosequencing was confirmed as a highly efficient deep-sequencing technique to characterize the bacterial diversity in drinking water. Sequences of *Alphaproteobacteria* and *Betaproteobacteria* dominated the bacterial community in drinking water with *Oxalobacteraceae* and *Methylobacteriaceae* as the most abundant bacterial families, which is consistent with the prominent abundance of these populations frequently detected in various freshwater environments where source waters originate. Bacterial populations represented by the most abundant sequences in drinking water were closely related to cultures of metabolically versatile bacterial taxa widely distributed in the environment, suggesting a potential link between environmental distribution, metabolic characteristics, and abundance in drinking water.

2.2. Introduction

The supply of safe drinking water is critical to public health. One of the most important challenges in drinking water supply is the control of microbial contamination. Despite the use of a suite of treatment processes including filtration and disinfection in water utilities, a small number of microorganisms remain present in treated water. Some of these microorganisms may be involved in microbially mediated corrosion and nitrification in water distribution systems, while others could be potentially pathogenic, presenting a poorly-understood public health risk (Berry et al. 2006). Data on waterborne disease outbreaks suggest that drinking water continues to be one of the most import media for infectious diseases worldwide (Ford 1999). Therefore, to develop effective control strategies and perform accurate risk assessment for microbial contamination in drinking water, it is necessary to understand the ecology of microorganisms in drinking water (Szewzyk et al. 2000).

Current practices in investigating the microbial ecology and biological integrity of drinking water have relied primarily on cultivation or 16S rRNA gene-based molecular techniques. Cultivation-based techniques, particularly heterotrophic plate count (HPC), have been used historically for the assessment of the microbiological quality of drinking water (Olson and Nagy 1984). However, the majority of microorganisms in nature are not yet culturable. To overcome this bias, cultivation-independent techniques, primarily molecular techniques such as clone library and fingerprinting analysis, have been used in the characterization of microbial populations in drinking water, revealing remarkable bacterial diversity and dynamics in drinking water through the analysis of 16S rRNA gene sequences (Eichler et al. 2006, Martiny et al. 2005, Poitelon et al. 2009, Revetta et al. 2010). However, limitations to the depth of DNA sequence coverage in conventional molecular techniques present a major challenge to studying the entire

dimension of microbial diversity using these methods, as population richness readily surpassing hundreds of phylotypes in drinking water (Poitelon et al. 2009).

The emergence of next-generation DNA sequencing technology, with its high-throughput DNA sequencing capacity (Huse et al. 2008), provides a unique opportunity to probe the potentially large bacterial diversity in drinking water. However, no application of high-throughput sequencing has been reported for the characterization of drinking water microbial communities. Therefore, in order to gain a more complete understanding of the ecology of microorganisms in drinking water, the objective of this study is to characterize the bacterial populations in drinking water. This study is the first application of pyrosequencing in drinking water, confirmed the validity of pyrosequencing as a valuable technique to characterize drinking water bacterial community composition with direct comparisons between pyrosequencing and conventional clone library analysis.

2.3. Materials and Methods

2.3.1. Water sample collection and handling

Two drinking water samples were used for the side-by-side comparison of pyrosequencing and clone library analysis in this study. Bulk drinking water samples were collected from the same faucet in the Environmental Engineering Laboratory on the campus of the University of Tennessee at Knoxville in December 2009 and June 2010, hereafter referred to as sample DWI and DWII, respectively. The water was supplied by a conventional water treatment plant with coagulation, rapid sand filtration, and chlorination. For each bulk water sampling, 150 L of tap water was collected and sealed in sterile polyethylene carboys following flushing the faucet for 10 minutes at maximum flow. An aliquot of the collected bulk water was taken for water quality

analysis. The water samples were subsequently dechlorinated by the addition of sodium thiosulfate to a final concentration of 50 mg/L as as previously described (Hill et al. 2007). Sodium polyphosphate was added to each water sample to a final concentration of 0.01% (w/v) as the dispersant for further processing.

2.3.2. Water quality analysis

Water quality analysis was performed for pH, turbidity, conductivity, free chlorine, dissolved organic carbon (DOC), sulfate, nitrate, and chloride. Turbidity was measured with a Hach 2100P turbidimeter (Hach Company, Loveland, Colorado, USA); conductivity was quantified with an Orion model 122 conductivity meter (Orion Research Inc., Boston, Massachusetts, USA); Free chlorine was measured following the "4500-Cl F" DPD Ferrous Titrimetric Method (APHA 2005); DOC was analyzed with a Shimadzu SSM-5000A TOC analyzer (Shimadzu Corporation, Kyoto, Japan); sulfate, nitrate, and chloride were quantified with a Dionex Ion Chromatograph ICS-2500 system (Dionex Corp., Sunnyvale, California, USA).

Table 2.1 Summary of water quality parameters

Water Sample	pH	Turbidity (NTU)	TOC (mg/L)	Conductivity (µS/cm)	Free Cl_2 (ppm)	Cl^- (ppm)	SO_4^{2-} (ppm)	NO_3^- (ppm)
DWI	6.8	0.07	1.1	280	2.2	11.0	17.4	1.7
DWII	6.8	0.10	1.3	260	2.1	10.9	16.9	1.3

2.3.3. Collection of microbial biomass by ultrafiltration

Microorganisms in drinking water were concentrated with a tangential-flow ultrafiltration system configured, prepared, and operated as previously described (Polaczyk et al. 2008). Briefly, all tubings and containers included in the ultrafiltration system were disinfected with 10% hypochlorous acid, washed with deionized water, and sterilized by autoclaving before use. The ultrafiltration system used sterile Fresenius Hemoflow F200NR polysulfone dialysis filters with a molecular weight cutoff of ~30,000 Daltons (Fresenius Medical Care, Waltham, Massachusetts, USA) as the ultrafilter. Pretreatment with 0.1% (w/v) sodium polyphosphate was performed to block the ultrafilters to minimize adhesion of microorganisms. Ultrafiltration was performed at a circulation rate of 2,000 mL/min and the final volume of the concentrate was ~300 mL following backflushing. The ultrafilters were discarded at the completion of each ultrafiltartion run to prevent cross-contamination.

Microbial biomass in the concentrate from ultrafiltration was further concentrated with Centricon Plus-70 centrifugal membrane filtration units with a molecular weight cutoff of ~30,000 Daltons (Millipore, Billerica, Massachusetts, USA) as previously described (Hill et al. 2007), followed by pelleting at 17,000 ×g at 4°C for 10 min. The pellets were immediately frozen at -80°C for subsequent nucleic acids extraction.

2.3.4. Pyrosequencing of bacterial populations in drinking water

Pyrosequencing of the 16S rRNA gene amplicon libraries was performed to characterize the bacterial diversity in drinking water following whole community DNA extraction and purification using a previously described method (Zhang et al. 2009). For each DNA sample, amplicon libraries were generated with primers targeting the V3 hypervariable region of the 16S rRNA gene: 338F (5'-ACTCCTACGGGAGGCAGCAG-3') and 533R (5'-

TTACCGCGGCTGCTGGCAC-3') (Huse et al. 2008). To pool multiple samples for one run of 454 sequencing, barcode sequences unique to each sample were attached to the primers following a previously described sample tagging approach (Hamady et al. 2008). Polymerase Chain Reaction (PCR) amplification was performed with the FastStart High Fidelity PCR system in a total volume of 50 μL, containing 5 μL of FastStart High Fidelity Reaction Buffer with 1.8 mM $MgCl_2$, 4% DMSO, 200μM dNTPs, 0.4 μM forward and reverse primers, 10-100 ng of DNA, and 2.5 U FastStart High Fidelity Enzyme Blend (Roche Diagnostics, Germany). The PCR reaction conditions were as follows: 1 cycle at 94 °C for 3 min, 20 cycles at 94 °C for 30 sec, 57 °C for 1 min, 72 °C for 2 min, and a final extension at 72 °C for 2 min. Amplicons were purified with the Qiagen PCR purification kit (Qiagen, Valencia, California, USA) and the Agencourt AMPure PCR purification system (Beckman Coulter, Danvers, Massachusetts, USA). The quality of each amplicon library was evaluated using the Agilent DNA 7500 kit with a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, USA). Equal molar quantities of amplicons from each water sample were pooled together. The pooled DNA was immobilized onto DNA capture beads and amplified through emulsion PCR using the GS FLX emPCR amplicon kit according to the manufacturer's protocols (454 Life Sciences, Branford, CT, USA). Sequencing of the PCR products was performed at the Center for Environmental Biotechnology at the University of Tennessee using a 454 Genome Sequencer FLX (454 Life Sciences, Branford, CT, USA).

2.3.5. Clone library analysis of bacterial populations in drinking water

Bacterial populations in drinking water were also surveyed with 16S rRNA gene-based clone library analysis. Clone libraries were constructed with community DNA from both the summer and winter samples using a previously described protocol with minor modification (Zhang et al.

2011). Briefly, bacterial 16S rRNA genes were amplified by bacterial universal primers 8F (5'-AGAGTTTGATCMTGGCTCAG-3') (Turner et al. 1999) and 907R (5'-CCGTCAATTCMTTTRAGTTT-3') (Lane 1991). Each PCR reaction mixture contained 0.4 μM of each primer, 200 μM dNTP, 2.5 U Ex Taq DNA polymerase, PCR buffer mix provided by the supplier of the Taq DNA polymerase (Takara, Madison, Wisconsin, USA), and 10 ng DNA template. PCR was performed with the following thermal cycling program: 94°C for 5 min, 9 cycles at 94°C for 1 min, 58°C for 1 min, and 72°C for 1 min, followed by 11 cycles at 94°C for 1 min, 54°C for 1 min, and 72°C for 1 min with a final extension at 72°C for 6 min. The amplified DNA products were purified with the Qiagen PCR purification kit (Qiagen, Valencia, California, USA) and cloned into pGEM-T Easy vector (Promega, Madison, Wisconsin, USA) according to the manufacturer's instructions. Clones were subsequently sequenced using M13 forward and reverse primers with an ABI 3730xl DNA Analyzer (Applied Biosystems, Foster city, CA, USA).

2.3.6. Analysis of pyrosequencing data

Sequences acquired by pyrosequencing were examined to remove low quality reads and sorted according to sample-specific barcodes using GS Amplicon Variant Analyzer software version 2.3 (Roche Diagnostics, Germany). Further removal of low quality reads and chimeric sequences was performed with the quality-filtering pipeline of the MOTHUR program (Schloss et al. 2009). The non-redundant sequences were aligned using the Needleman-Wunsch and NAST algorithms implemented in MOTHUR against the SILVA database alignment (http://www.arb-silva.de/). A pairwise distance matrix was constructed for all subsequent microbial community ecological analyses, including community diversity and rarefaction analysis. Operational taxonomic units (OTUs) were assigned with the average neighbor clustering algorithm and taxonomy was

assigned using the naïve Bayesian rRNA classifier of Ribosomal Database Project (RDP) with a bootstrap cutoff of 80% (Wang et al. 2007).

2.3.7. Analysis of clone library results

Partial 16S rRNA gene sequences from the clone libraries were first assembled using Sequencher 4.9 (Gene Codes, Ann Arbor, Michigan, USA). The assembled sequences were analyzed similarly as described in 'Analysis of pyrosequencing data', with the exception that sequences shorter than 700 bp were excluded from analysis due to the longer sequence reads in clone library analysis. Additionally, select 16S rRNA gene sequences from the clone libraries were searched for homologues using the BLAST program at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST/), aligned with homologous sequences using ClustalX (Thompson et al. 1997), and used for the construction of phylogenetic trees with the neighbor-joining algorithm (1,000 bootstrap re-samplings) using MEGA 4.0 (Tamura et al. 2007). The non-redundant 16S rRNA gene sequences obtained from clone library analysis were deposited at GenBank under the following accession numbers HQ711889 to HQ711923.

2.4. Results and Discussion

2.4.1. Community coverage by pyrosequencing and clone library

Pyrosequencing of the 16S rRNA gene amplicons from the summer and winter drinking water samples yielded 6,419 and 4,829 high quality reads for DWI and DWII, respectively. These sequences had an average length of 144 bp, representing 1,463 unique sequences. In comparison, clone library analysis as a more conventional culture-independent technique performed in

parallel with pyrosequencing yielded a total of 173 16S rRNA gene sequences averaging 873 bp in length, representing 35 unique 16S rRNA gene sequences. Overall, the number of 16S rRNA gene sequences recovered by pyrosequencing (11,248) was much greater than that obtained from clone library analysis (173) in this study. Similarly, the number of sequences from pyrosequencing is also much greater than those reported in other studies using 16S rRNA gene-based conventional molecular analyses of drinking water, which are typically in the hundreds at the most (Poitelon et al. 2009, Revetta et al. 2010), suggesting the utility of pyrosequencing for assessing the full extent of microbial diversity in drinking water.

2.4.2. Comparison of community compositions revealed by pyrosequencing and clone library

Since abundant populations are typically of the most interest, comparisons were first made between the dominant bacterial populations identified by pyrosequencing and clone library. A large majority (94.8%) of sequences identified by pyrosequencing represented *Proteobacteria* (Figure 2.1), followed by *Actinobacteria* as the distant second in abundance with *Firmicutes* rounding out the three phyla with a relative abundance greater than 1%. The same distribution was also found by clone library analysis with *Proteobacteria* dominating the bacterial community followed by *Actinobacteria* and *Firmicutes* as the only phyla represented by > 1% of the sequences (Figure 2.1). Thus, the same bacterial populations dominating the drinking water samples were identified by both clone library and pyrosequencing, demonstrating the consistency between these two methods.

More detailed comparisons of bacterial community composition were made between pyrosequencing and clone library at the family level. Both methods were consistent in the identification of *Methylobacteriaceae* and *Oxalobacteraceae* as the dominant taxa in DWI and DWII, respectively (Figure 2.2). Clone library analysis was also able to confirm the detection of

most of the less abundant bacterial families identified by pyrosequencing; however, clone library failed to recover sequences associated with *Acetobacteraceae* in DWI and *Methylobacteriaceae* in DWII, both of which were detected by pyrosequencing (Figure 2.2). Overall, all taxa identified by clone library were also detected by pyrosequencing; however, less abundant taxa identified by pyrosequencing could be missed by clone library, thus highlighting the ability of pyrosequencing to detect populations with low abundance. Nevertheless, findings from pyrosequencing and clone library analysis were highly consistent, supporting the validity and utility of high-throughput sequencing for studying microorganisms in drinking water. It should be noted that the HPC counts of the two water samples were similar, both in the range of 0 - 10 CFU/100 mL, further supporting the advantage of cultivation-independent techniques over cultivation-dependent techniques in providing a more compete profile of the microbial community in drinking water.

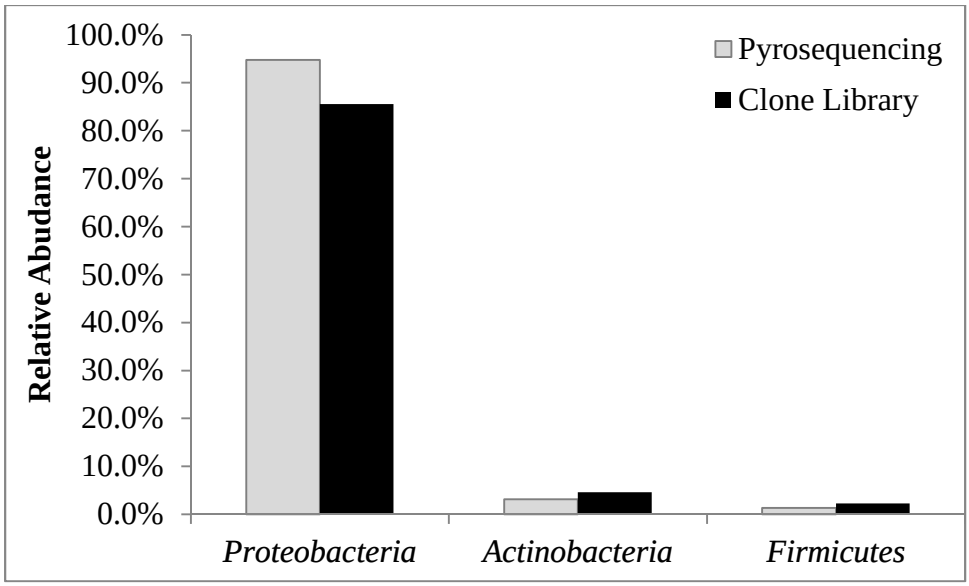

Figure 2.1 Dominant bacterial populations with relative sequence abundance >1% as identified by pyrosequencing and clone library analysis. Shown are bacterial populations classified at the phylum level based on 16S rRNA gene sequences from water samples DWI and DWII.

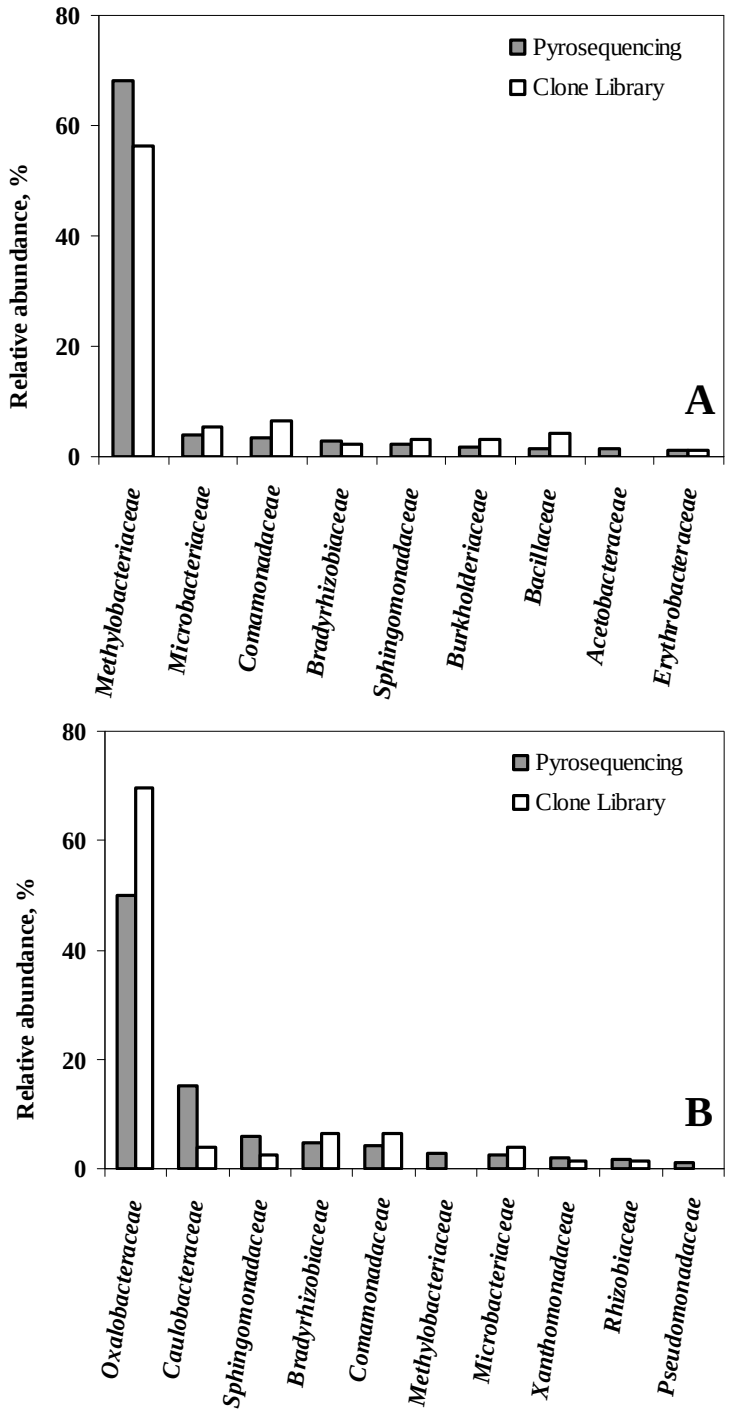

Figure 2.2 Comparison of bacterial community compositions identified by pyrosequencing and clone library at the family level in drinking water: (A) winter; and (B) summer. Shown are bacterial taxa identified at the family level with relative sequence abundance greater than 1%.

2.4.3. Characteristics of bacterial community composition

Given the greater depth of coverage provided by pyrosequencing, the 16S rRNA gene sequences recovered from pyrosequencing were used for the characterization of bacterial community in the drinking water samples. These sequences were distributed into six bacterial phyla, with the majority (94.8%) representing Proteobacteria (Figure 2.3A). Actinobacteria was the distant second in abundance, accounting for 3.2% of the sequences. Firmicutes, represented by 1.4% of the sequences, was the third phylum with a relative abundance greater than 1%. The other phyla, *Acidobacteria*, *Bacteroidetes*, and *Deinococcus-Thermus*, in total accounted for 0.74% of the sequences.

Dominating the bacterial community in drinking water, *Proteobacteria* was found to be represented primarily by *Alphaproteobacteria* and *Betaproteobacteria*, together making up 96.2% of the *Proteobacteria* sequences (Figure 2.3B). Representative genera identified by pyrosequencing included *Methylobacterium*, *Caulobacter*, and *Sphingomonas* in *Alphaproteobacteria*, and *Massilia* and *Acidovorax* in *Betaproteobacteria*. Sequences of *Gammaproteobacteria* and *Deltaproteobacteria* were also present, albeit as very minor constituents, comprising 3.3 and 0.1% of the *Proteobacteria* community, respectively (Figure 2.3B).

The dominance of *Alphaproteobacteria* and *Betaproteobacteria* observed in this study (Figure 2.3B) is consistent with previous reports on the bacterial diversity of drinking water (Eichler et al. 2006, Poitelon et al. 2010). Interestingly *Alphaproteobacteria* and *Betaproteobacteria* were also found to dominate the *Proteobacteria* populations in many freshwater habitats (Zwart et al. 2002). Since surface freshwater was used as the source water for drinking water treatment in these studies, the similar pattern of *Proteobacteria* dominance points

to a potential link between source water and drinking water. Indeed, evidence from a study comparing the bacterial community compositions in source water and finished water suggests that the microbial community in source water had significant influence on that of the drinking water (Eichler et al. 2006). However, *Alphaproteobacteria* and *Betaproteobacteria* have also been found to be the primary constituents of biofilm in some water distribution systems (Hong et al. 2010, Mathieu et al. 2009, Yu et al. 2010), presenting the possibility that biofilm might also serve as a source of microorganisms to drinking water.

Analysis of the bacterial community composition revealed considerable temporal changes in bacterial community composition in drinking water (Figure 2.2). The contribution of *Methylobacteriaceae* to the bacterial community diminished from 68.2% in winter (DWI) as the most numerous representative of *Alphaproteobacteria* to a mere 2.9% in the summer (DWII). *Caulobacteraceae*, another representative of *Alphaproteobacteria*, however, emerged as a minor population (0.9% abundance) in winter (DWI) to become the second most abundant taxon in summer, covering 15.2% of the sequences. The quantities of two other abundant members of *Alphaproteobacteria*, *Sphingomonadaceae* and *Bradyrhizobiaceae*, remained relatively stable. In contrast, the changes in the *Betaproteobateria* could be solely attributed to *Oxalobacteraceae*, which was elevated to the dominant bacterial family in summer from near non-existence in winter (Figure 2.2). These results reveal that bacterial community composition in drinking water might be subjected to significant seasonal changes. Due to the limited scope of this study, a complete temporal pattern of drinking water microbial community was not obtained. However, the dynamics of microbial communities in drinking water is potentially important and warrants further investigation with more frequent sampling throughout the year.

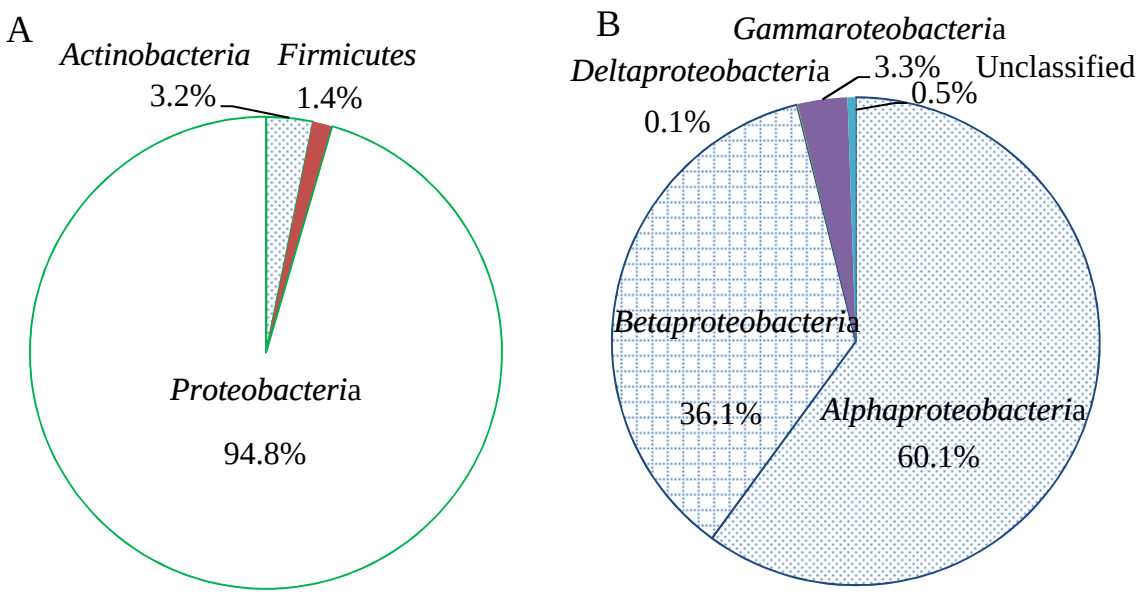

Figure 2.3 Bacterial composition of drinking water as revealed by pyrosequencing of 16S rRNA genes in (A) the whole community and (B) the phylum of *Proteobacteria*. Bacterial taxa with relative sequence abundance greater than 1% are labeled, with the percent values showing the relative abundance in all sequence reads obtained from the winter and summer samples.

2.4.4. Phylogenetic analysis of representative bacterial phylotypes in drinking water

A number of sequences representing bacterial families with high relative abundance in this study were closely related to known bacteria cultures (Figure 2.4). Sequences of *Methylobacteriaceae*, which was the most represented taxon at the family level in the winter water sample DWI (Figure 2.2A), shared >99% identity with the 16S rRNA sequence of *Methylobacterium radiotolerans*. *Methylobacterium* spp. are widely distributed in aquatic environments and have been isolated from drinking water and biofilm in water distribution system (Hiraishi et al. 1995). While the physiology of *M. radiotoleras* is not well characterized, recent studies on *Methylobacterium* isolates from drinking water have shown that these bacteria are capable of forming biofilm and utilizing a diverse group of carbon substrates in addition to C1 compounds (Gallego et al. 2005, 2006a, Simoes et al. 2010).

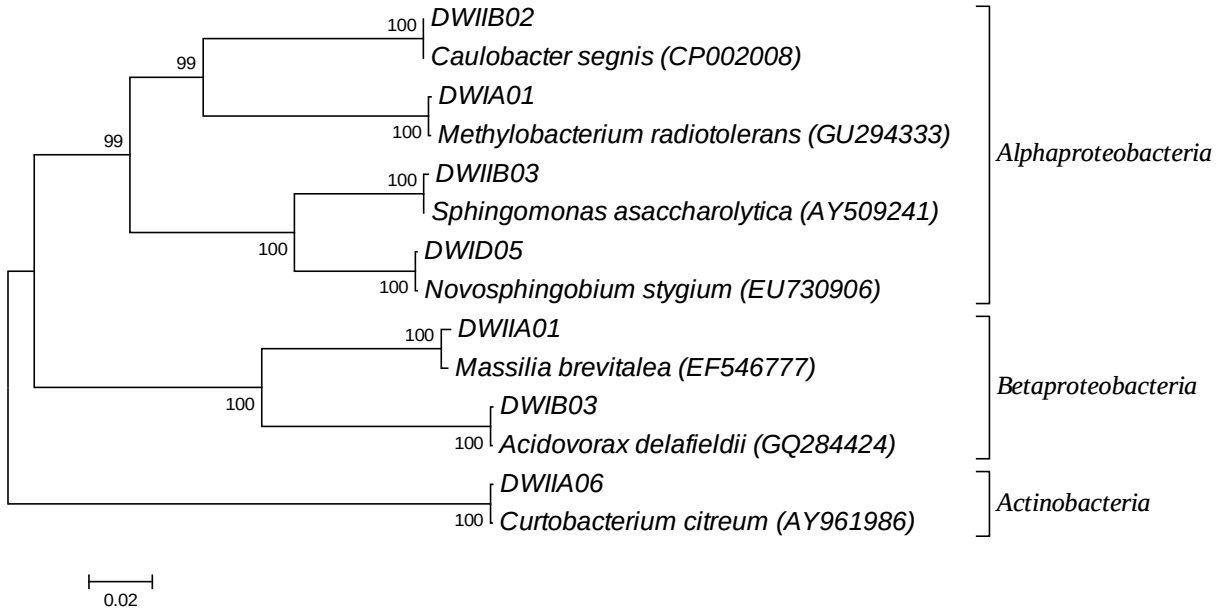

Figure 2.4 Neighbor-joining phylogenetic tree showing relationships of representative partial 16S rRNA gene sequences cloned from drinking water to close relatives. Clones from this study are in bold .The scale bar represents the number of substitutions per sequence position.

Sequences representing *Oxalobacteraceae*, which was dominant in DWII (Figure 2.2B), was closely related to a soil isolate *Massilia brevitalea*, capable of using a variety of organic compounds including volatile fatty acids (Zul et al. 2008). Interestingly, a drinking water isolate of *Massilia* exhibited the tendency to form pellicles in liquid medium, suggesting the potential of these bacteria in biofilm formation (Gallego et al. 2006b). Sequences of *Caulobacteraceae*, the 3rd most represented bacterial family in this study, were closely related to *Caulobacter segnis*, which is a member of the freshwater *Caulobacter* cluster ubiquitous in oligotrophic environments (Abraham et al. 1999). Additionally, *Caulobacter* is well known to be capable of surface attachment and biofilm formation (Entcheva-Dimitrov and Spormann 2004), providing another advantage facilitating the existence of these bacteria in drinking water systems.

Sequences recovered in this study were also found to share >99% identity with those of two *Sphingomonads*, *Novosphingobium stygium* and *Sphingomonas asaccharolytica* (Figure.

2.4). While isolated from different environments, i.e. subsurface sediment and plant roots, respectively, these two bacteria are members of the metabolically diverse *Sphingomonadaceae* ubiquitous in various environments (Takeuchi et al. 1995). Similarly, *Acidovorax delafieldii*, a soil bacterium closely related to sequences classified as *Comamodadaceae* in this study, is versatile in substrate utilization and energy metabolism (Willems et al. 1990). Sequences of *Microbacteriaceae,* another bacterial family well represented in the drinking water bacterial community, were >99% similar to the 16S rRNA gene sequence of *Curtobacterium citreum*, which is capable of utilizing various sugars and organic acids (Yamada and Komagata 1972).

In conclusion, bacterial populations highly represented in drinking water appear to share common features of broad distribution in the environment, versatility in substrate utilization, and potential in bio film formation. Thus, the observation that quantitatively important bacterial populations in drinking water were frequently related to members of metabolically versatile bacterial taxa widely distributed in the environment suggests a potential link between environmental distribution, metabolic characteristics, and abundance in drinking water. Since source water is the most relevant entry point for microorganisms in the environment, the microbiological quality of drinking water could be impacted by source water management practices. However, the contribution of source water to the microbial community in finished water needs to be further evaluated by comparing the bacterial communities between source water and finished water in various drinking water treatment systems.

2.5. Conclusions

This study represents the first application of high-throughput sequencing in drinking water microbiology in the literature. Validated by clone library analysis, pyrosequencing was con-

firmed as a highly efficient deep-sequencing technique to characterize the bacterial diversity in drinking water. Sequences of Alphaproteobacteria and Betaproteobacteria dominated the bacterial community in drinking water, which is consistent with the prominent abundance of these populations frequently detected in various freshwater environments where source waters originate. Bacterial populations represented by the most abundant sequences in drinking water were closely related to cultures including *Methylobacterium radiotolerans*, *Massilia brevitalea*, and *Caulobacter segnis*, which were shown to be frequently related to members of metabolically versatile bacterial taxa widely distributed in the environment, suggesting a potential link between environmental distribution, metabolic characteristics, and abundance in drinking water. Further, the revelation of significant temporal dynamics in microbial community composition suggests that a more comprehensive understanding of the microbial populations existing in drinking water requires more frequent sampling with cultivation-independent microbial ecology techniques. More importantly, the identification of sequences unrelated to common microbial indicators demonstrates the presence of yet-to-be-characterized disinfectant-resistant microbial risks in drinking water, suggesting the importance of developing additional microbial indicators and techniques for the monitoring of these microbial risks.

2.6. References

Abraham, W.-R., Strompl, C., Meyer, H., Lindholst, S., Moore, E.R.B., Christ, R., Vancanneyt, M., Tindall, B.J., Bennasar, A., Smit, J. and Tesar, M. (1999) Phylogeny and polyphasic taxonomy of *Caulobacter* species. Proposal of *Maricaulis* gen. nov. with *Maricaulis maris* (Poindexter) comb. nov. as the type species, and emended description of the genera *Brevundimonas* and *Caulobacter*. International Journal of Systematic Bacteriology 49(3), 1053-1073.

APHA (2005) Standard methods for the examination of water and wastewater, American Public Health Association, Washington, D.C.

Berry, D., Xi, C.W. and Raskin, L. (2006) Microbial ecology of drinking water distribution systems. Current Opinion in Biotechnology 17(3), 297-302.

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.

Entcheva-Dimitrov, P. and Spormann, A.M. (2004) Dynamics and control of biofilms of the oligotrophic bacterium Caulobacter crescentus. Journal of Bacteriology 186(24), 8254-8266.

Ford, T.E. (1999) Microbiological safety of drinking water: United States and global perspectives. Environmental Health Perspectives 107, 191-206.

Gallego, V., García, M.T. and Ventosa, A. (2005) *Methylobacterium hispanicum* sp. nov. and *Methylobacterium aquaticum* sp. nov., isolated from drinking water. International Journal of Systematic and Evolutionary Microbiology 55(1), 281-287.

Gallego, V., García, M.T. and Ventosa, A. (2006a) *Methylobacterium adhaesivum* sp. nov., a methylotrophic bacterium isolated from drinking water. International Journal of Systematic and Evolutionary Microbiology 56(2), 339-342.

Gallego, V., Sánchez-Porro, C., García, M.T. and Ventosa, A. (2006b) *Massilia aurea* sp. nov., isolated from drinking water. International Journal of Systematic and Evolutionary Microbiology 56(10), 2449-2453.

Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J. and Knight, R. (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nature Methods 5(3), 235-237.

Hill, V.R., Kahler, A.M., Jothikumar, N., Johnson, T.B., Hahn, D. and Cromeans, T.L. (2007) Multistate evaluation of an ultrafiltration-based procedure for simultaneous recovery of enteric microbes in 100-liter tap water samples (vol 73, pg 4218, 2007). Applied and Environmental Microbiology 73(19), 6327-6327.

Hiraishi, A., Furuhata, K., Matsumoto, A., Koike, K.A., Fukuyama, M. and Tabuchi, K. (1995) Phenotypic and genetic diversity of chlorine-resistant *Methylobacterium* strains isolated from various environments. Applied and Environmental Microbiology 61(6), 2099-2107.

Hong, P.Y., Hwang, C.C., Ling, F.Q., Andersen, G.L., LeChevallier, M.W. and Liu, W.T. (2010) Pyrosequencing Analysis of Bacterial Biofilm Communities in Water Meters of a Drinking Water Distribution System. Applied and Environmental Microbiology 76(16), 5631-5635.

Huse, S.M., Dethlefsen, L., Huber, J.A., Welch, D.M., Relman, D.A. and Sogin, M.L. (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. Plos Genetics 4(11), 10.

Lane, D.J. (1991) Nucleic acid techniques in bacterial systematics. Stackebrandt, E. and Goodfellow, M. (eds), pp. 115-175 John Wiley & Sons, New York, N.Y.

Martiny, A.C., Albrechtsen, H.J., Arvin, E. and Molin, S. (2005) Identification of bacteria in biofilm and bulk water samples from a nonchlorinated model drinking water distribution system: Detection of a large nitrite-oxidizing population associated with Nitrospira spp. Applied and Environmental Microbiology 71(12), 8611-8617.

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.

Olson, B.H. and Nagy, L.A. (1984) Microbiology of potable water. Advances in Applied Microbiology 30, 73-132.

Poitelon, J.-B., Joyeux, M., Welté, B., Duguet, J.-P., Prestel, E., Lespinet, O. and DuBow, M.S. (2009) Assessment of phylogenetic diversity of bacterial microflora in drinking water using serial analysis of ribosomal sequence tags. Water Research 43(17), 4197-4206.

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 & Biotechnology 37(2), 117-128.

Polaczyk, A.L., Narayanan, J., Cromeans, T.L., Hahn, D., Roberts, J.M., Amburgey, J.E. and Hill, V.R. (2008) Ultrafiltration-based techniques for rapid and simultaneous concentration of multiple microbe classes from 100-L tap water samples. Journal of Microbiological Methods 73(2), 92-99.

Revetta, R.P., Pemberton, A., Lamendella, R., Iker, B. and Santo Domingo, J.W. (2010) Identification of bacterial populations in drinking water using 16S rRNA-based sequence analyses. Water Research 44(5), 1353-1360.

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., Stres, B., Thallinger, G.G., Van Horn, D.J. and Weber, C.F. (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology 75(23), 7537-7541.

Simoes, L.C., Simoes, M. and Vieira, M.J. (2010) Adhesion and biofilm formation on polystyrene by drinking water-isolated bacteria. Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology 98(3), 317-329.

Szewzyk, U., Szewzyk, R., Manz, W. and Schleifer, K.H. (2000) Microbiological safety of drinking water. Annual Review of Microbiology 54, 81-127.

Takeuchi, M., Sakane, T., Yanagi, M., Yamasato, K., Hamana, K. and Yokota, A. (1995) Taxonomic Study of Bacteria Isolated from Plants: Proposal of Sphingomonas rosa sp. nov., Sphingomonas pruni sp. nov., Sphingomonas asaccharolytica sp. nov., and Sphingomonas mali sp. nov. . International Journal of Systematic Bacteriology 45(2), 334-341.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24(8), 1596-1599.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25(24), 4876-4882.

Turner, S., Pryer, K.M., Miao, V.P.W. and Palmer, J.D. (1999) Investigating deep phylogenetic relationships among cyanobacteria and plastids by small submit rRNA sequence analysis. Journal of Eukaryotic Microbiology 46(4), 327-338.

Wang, Q., Garrity, G.M., Tiedje, J.M. and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73(16), 5261-5267.

Willems, A., Falsen, E., Pot, B., Jantzen, E., Hoste, B., Vandamme, P., Gillis, M., Kersters, K. and De Ley, J. (1990) *Acidovorax*, a new genus for *Pseudomonas facilis*, *Pseudomonas delafieldii*, E. Falsen (EF) group 13, EF group 16, and several clinical isolates, with the species *Acidovorax facilis* comb. nov., *Acidovorax delafieldii* comb. nov., and *Acidovorax temperans* sp. nov. International Journal of Systematic Bacteriology 40(4), 384-398.

Yamada, K. and Komagata, K. (1972) Taxonomic studies on coryneform bacteria V. Classification of coryneform bacteria. Journal of General and Applied Microbiology 18(6), 417-431.

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.

Zhang, Y., Zhang, X.L., Zhang, H.W., He, Q., Zhou, Q.X., Su, Z.C. and Zhang, C.G. (2009) Responses of soil bacteria to long-term and short-term cadmium stress as revealed by microbial community analysis. Bulletin of Environmental Contamination and Toxicology 82(3), 367-372.

Zhang, Y., Zamudio Cañas, E.M., Zhu, Z.W., Linville, J.L., Chen, S. and He, Q. (2011) Robustness of archaeal populations in anaerobic co-digestion of dairy and poultry wastes. Bioresource Technology 102(2), 779-785.

Zul, D., Wanner, G. and Overmann, J. (2008) *Massilia brevitalea* sp. nov., a novel betaproteobacterium isolated from lysimeter soil. International Journal of Systematic and Evolutionary Microbiology 58(5), 1245-1251.

Zwart, G., Crump, B.C., Kamst-van Agterveld, M.P., Hagen, F. and Han, S.-K. (2002) Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. Aquatic Microbial Ecology 28(2), 141-155.

Chapter 3. Variations in Drinking Water Bacterial Community Composition

3.1. Abstract

The understanding of the microbial assembly mechanisms in drinking water is still rudimentary. Bacterial community composition in drinking water were investigated using both pyrosequnencing and clone library analysis. A broad range of diverse bacteria predominated by *proteobacteria* were identified. The dominant populations persistent in all the drinking water were *Alphaproteobacteria*, *Betaproteobacteria*, and *Actinobacteria*. Several core abundant families were detected in all the water samples: *Sphingomonadaceae, Caulobacteraceae, Methylobacteriaceae, Oxalobacteraceae*, *Comamonadaceae*, *Mycobacteriaceae* and *Peptostreptococcaceae*, which represented by *Sphingomonas*, *Novosphingobium*, *Caulobacter*, *Methylobacterium*, *Massilia*, *Acidovorax*, and *Mycobacterium*. Principal component analysis and cluster analysis showed that bacterial community compositions were influenced by source water and environmental variables.

3.2. Introduction

Despite strict regulation and frequent monitoring for drinking water, bacteria can be persistent in drinking water. Some of these bacteria may be involved in microbially mediated corrosion and nitrification in water distribution systems, while others could be potentially pathogens. The regrowth of bacteria in drinking water distribution system (DWDS) will deteriorate water quality and directly threaten public health as the outbreak of waterborne infectious diseases (Craun et al. 2010, Jjemba et al. 2010, Li et al. 2010). Therefore, it is necessary to understand the ecology of microorganisms in drinking water, perform accurate risk assessment, and develop efficient strategies to control the microbial contamination in drinking water.

In recent years cultured and uncultured molecular approaches have been used to study the bacterial community in DWDS. Metagenomic surveys based on 16S rRNA phylogenetic analysis showed that *Alphaproteobacteria*, *Betaproteobacteria*, and *Gammaproteobacteria* were the most abundant groups inhabiting in DWDS followed by *Firmicutes* and *Actinobacteria* (Eichler et al. 2006, Hong et al. 2010, Norton and LeChevallier 2000). Some researchers have interest in detecting pathogens others targeted on bacteria with ecological functions such as nitrite, ammonia oxidizer and disinfection byproducts degrader (Jjemba et al. 2010, Leach et al. 2009, Lipponen et al. 2004, Martiny et al. 2005, Regan et al. 2003, Zhang et al. 2009). These studies depend on clone library, denaturing gradient gel electrophoresis (DGGE) or terminal restriction length polymorphism (T-RFLP), which may underestimate the actual number and diversity due to the limited separating resolution and sequencing size. Therefore they may highlight the predominant groups and neglect the low abundance groups which may possess ecological functions or potential pathogens. The emergence of next-generation DNA sequencing dramatically increased sequencing capacity and sampling coverage, which allows a much higher throughput than the previous molecular techniques. Pyrosequencing of partial 16S rRNA genes have been employed in various ecosystems such as soil, biofilm and human body (Costello et al. 2009, Hong et al. 2010, Will et al. 2010). And enough sampling effort was reported in these studies.

Microfiltration with membranes is considered an attractive alternative to traditional sand filtration due to the ability of remove microorganisms as well as particles by sieving water through smaller pores (0.2 μm). However, the effect of membrane filtration process on bacterial diversity in drinking water is still rudimentary. Previous research showed that bacterial communities in drinking water can be affected by source water and treatment process (Eichler et

al. 2006, Poitelon et al. 2010, Stewart et al. 1990). We sampled five tap water supplied by five water treatment plants in Knox County. One of them changed to membrane filtration and the other four still using traditional sand filtration. The objective of this study is to compare the bacterial community in consumers' tap supplied by different water treatment plants, characterize the bacterial community composition using high throughput pyrosequencing technology, and explain the impaction of environmental variables on the community composition.

3.3. Material and Methods

3.3.1. Study sites and water sampling

Study sites covered five water treatment plants serving areas in Knox County Tennessee USA (Figure 3.1). Two separate river branches provide source water for these water treatment plants. Water treatment plant WTP-1F, WTP-2R and WTP-3S use surface water from Tennessee River as source water. WTP-4L and WTP-5P treat surface water from Clinch River. The first four water treatment plants applied traditional treatment process including alum coagulation, flocculation, sedimentation, sand filtration, and disinfection with chlorine. The raw water in WTP-5P is treated through a membrane filtration process. After coagulation and flocculation, the raw water is pumped through submerged microfiltration Siemens membranes and followed by a disinfection step. Tap water from five locations in Knox County affiliated to five water treatment plants were sampled in June 2010 as shown in Figure 3.1. Before sampling cold tap water faucet was flushed for 10 min. 150 liters of water at each sampling sites were collected with autoclaved polyethylene carboys and transported into laboratory. An aliquot of the collected bulk water was taken for water quality analysis. The rest water samples were used for bacteria collection.

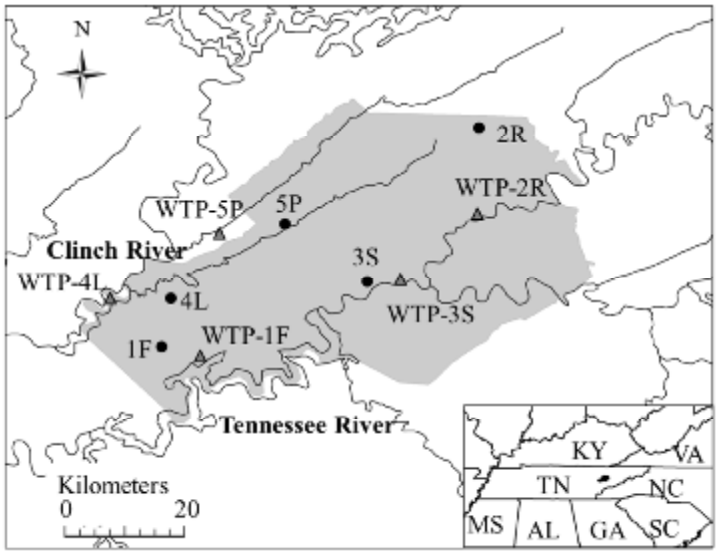

Figure 3.1 Drinking water sampling sites in Knox County Tennessee USA. The grey area is Knox County. Black dots, grey triangles and lines indicate sampling sties, water treatment plants and streams respectively.

3.3.2. Water quality analysis

All water samples were characterized using the following water quality parameters: pH, turbidity, conductivity, free chlorine, dissolved organic carbon (DOC), sulfate, nitrate, and chloride. Turbidity was measured with a Hach 2100N turbidimeter (Hach Company, Loveland, Colorado, USA); conductivity was quantified with an Orion model 122 conductivity meter (Orion Research Inc., Boston, Massachusetts, USA); Free chlorine was measured following standard method "4500-Cl F" DPD Ferrous Titrimetric Method (APHA 2005); DOC was analyzed with a Shimadzu SSM-5000A TOC analyzer (Shimadzu Corporation, Kyoto, Japan); sulfate, nitrate, and chloride were quantified with a Dionex Ion Chromatograph ICS-2500 system (Dionex Corp., Sunnyvale, California, USA). Heterotrophic plate counts (HPC) were measure using R2A agar at 28°C as previously described (Reasoner and Geldreich 1985).

3.3.3. Bacteria collection by ultrafiltration

Bacteria in drinking water were concentrated with a tangential-flow ultrafiltration system configured, prepared, and operated as previously described (Polaczyk et al. 2008). Briefly, all tubings and containers included in the ultrafiltration system were disinfected with 10% hypochlorous acid, washed with deionized water, and sterilized by autoclaving before use. The water samples were dechlorinated by the addition of sodium thiosulfate to a final concentration of 50 mg/L and sodium polyphosphate was added to each water sample to a final concentration of 0.01% (w/v) as the dispersant as previously described (Hill et al. 2007, Mull and Hill 2009, Polaczyk et al. 2008).

Water sample was immediately ultrafiltered to about 300 ml through hollow-fiber Fresenius Hemoflow F200NR polysulfone dialysis filters with a molecular weight cutoff of 30 kDa (Fresenius Medical Care, Waltham, MA). The concentrated water sample was brought to a final volume of 2 ml using 30 kDa Centricon Plus-70 units (Millipore, Billerica, MA) according to the manufacturer's instructions. Bacteria pellet was collected after centrifuging 2 ml of concentrate at 17,000 g at 4°C for 10 min (accuSpin Micro 17R, Thermo Fisher Scientific) and was immediately stored at -80°C for DNA extraction.

3.3.4. Pyrosequencing of 16S rDNA amplicons

The whole genome DNA was extracted using FastDNA spin kit for soil (MP Biomedicals, Santa Anna, CA). Pyrosequencing of the 16S rRNA gene amplicon libraries was performed to characterize the bacterial diversity in drinking water. For each DNA sample, amplicon libraries were generated with primers targeting the V3 hypervariable region of the 16S rRNA gene: 338F (5’-ACTCCTACGGGAGGCAGCAG-3’) and 533R (5’-TTACCGCGGCTGCTGGCAC-3’) (Huse et al. 2008). Barcode sequences unique to each sample were attached to both primers

following a previously described sample tagging approach (Hamady et al. 2008). Polymerase Chain Reaction (PCR) amplification was performed with the FastStart High Fidelity PCR system in a total volume of 50 μL, containing 5 μL of FastStart High Fidelity Reaction Buffer with 1.8 mM $MgCl_2$, 200μM dNTPs, 0.4 μM forward and reverse primers, 10-100 ng of DNA, and 2.5 U FastStart High Fidelity Enzyme Blend (Roche Diagnostics, Germany). PCR was performed with the following thermal cycling program: 1 cycle at 94 °C for 3 min, 20 cycles at 94 °C for 30 sec, 57 °C for 1 min, 72 °C for 2 min, and a final extension at 72 °C for 2 min. Amplicons were purified with the Qiagen PCR purification kit (Qiagen, Valencia, California, USA) and the Agencourt AMPure PCR purification system (Beckman Coulter, Danvers, Massachusetts, USA). The quality of each amplicon library was evaluated using the Agilent DNA 7500 kit with a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, USA). Equal molar quantities of amplicons from each water sample were pooled together. The pooled DNA was immobilized onto DNA capture beads and amplified through emulsion PCR using the GS FLX emPCR amplicon kit according to the manufacturer's protocols (454 Life Sciences, Branford, CT, USA). Sequencing of the PCR products was performed at the Center for Environmental Biotechnology at the University of Tennessee using a 454 Genome Sequencer FLX (454 Life Sciences, Branford, CT, USA).

3.3.5. Clone library analysis of 16S rDNA genes

16S rRNA gene-based clone libraries were also constructed to analyze the bacterial community composition in drinking water. Clone libraries were from five water samples were constructed with community DNA using a previously described protocol with minor modification (Zhang et al. 2011). Briefly, bacterial 16S rRNA genes were amplified by bacterial universal primers 8F (5'-AGAGTTTGATCMTGGCTCAG-3') and 907R (5'-CCGTCAATTCMTTTRAGTTT-3')

(Lane 1990, Martin-Laurent et al. 2001). PCR reaction mixture contained PCR buffer mix provided by the supplier of Taq DNA polymerase, 0.4 μM of each primer, 200 μM dNTP, 2.5 U Ex Taq DNA polymerase (Takara, Madison, Wisconsin, USA), and 10 ng DNA template. PCR was performed with the following thermal cycling program: 94°C for 5 min, 9 cycles at 94°C for 1 min, 58°C for 1 min, and 72°C for 1 min, followed by 11 cycles at 94°C for 1 min, 54°C for 1 min, and 72°C for 1 min with a final extension at 72°C for 6 min. The amplified DNA products were purified using Qiagen PCR purification kit (Qiagen, Valencia, California, USA) and cloned into pGEM-T Easy vector (Promega, Madison, Wisconsin, USA) according to the manufacturer's instructions. Clones were subsequently sequenced using M13 forward and reverse primers with an ABI 3730xl DNA Analyzer (Applied Biosystems, Foster city, CA, USA).

3.3.6. Sequence analysis

Sequences acquired by pyrosequencing were parsed and trimmed according to sample-specific barcodes using GS Amplicon Variant Analyzer software version 2.3 (Roche Diagnostics, Germany). The analysis of 16S rRNA gene sequences obtained from pyrosequencing and clone library were both performed with MOTHUR program version 1.21.0 (Schloss et al. 2009). After the removal of short sequences (pyrosequencing <100 bp, clone sequences < 600 bp) the dataset was de-replicated to eliminate duplicate sequences. After checking and removing chimera sequences, the remaining sequences were de-replicated again and low-quality sequences were removed as described previously (Huse et al. 2010). The non-redundant sequences were aligned using the Needleman-Wunsch and NAST algorithms implemented in MOTHUR against the SILVA database alignment (http://www.arb-silva.de/). A pairwise distance matrix was constructed for all subsequent microbial community ecological analyses, including community diversity and rarefaction analysis. Operational taxonomic units (OTUs) were assigned with the

average neighbor clustering algorithm and taxonomy was assigned using the naïve Bayesian rRNA classifier of Ribosomal Database Project (RDP) with a bootstrap cutoff of 80% (Wang et al. 2007). Multiple sample analysis was performed using Unifrac dendrograms. Phylogenetic trees for clone library sequences were constructed by aligning with their closest relatives in NCBI database (http://www.ncbi.nlm.nih.gov/BLAST/) through ClustalX1.83 and clustering with neighbor-joining algorithm (1000 bootstrap re-samplings) through Mega 5.0 (Tamura et al. 2011, Thompson et al. 1997).

3.3.7. Statistical analysis

In order to examine the relationship between water quality parameters and bacterial community composition principal component analysis (PCA) was performed by using CANOCO 4.5 (Microcomputer Power, Ithaca, NY). To remove the weight of rare species, a relative abundance above 2% at family level were used. Water quality variable: Turbidity, conductivity, hardness, DOC and free chorine which reflect drinking water properties were selected for PCA.

Nucleotide sequence accession numbers

The 16S rRNA nucleotide sequences from this study were deposited in the GenBank database under the accession numbers JF460939 to JF461001. 454 pyrosequencing raw data was submitted to NCBI GenBank Short Read Archive (SRA) with the accession number SRA036912.1.

3.4. Results and Discussion

3.4.1. Taxonomic composition of bacterial communities in drinking water

In order to characterize bacteria community composition in drinking water, five tap waters at customer end represented five water treatment plants serving areas in Knox County were sampled. Bacterial communities were analyzed by 16S rDNA based high throughput pyrosequencing and clone library Sanger sequencing. A total of 32,039 reads with an average length of ~ 150 bp were generated from pyrosequencing. At 80% cut off level, 98.2% of those sequences can be classified to bacterial phyla. Total 10 bacterial phyla were identified: *Proteobacteria* (85.5%), *Actinobacteria* (9.3%), *Firmicutes* (1.9%), *Planctomycetes* (0.6%), *Bacteroidetes* (0.5%), *Spirochaetes* (0.1%), *Acidobacteria* (0.1%) and 0.1% other rare phyla. *Proteobacteria* was found predominate in all water samples with relative abundance above 95% in sample 1F, 2R and 3S, and above 72% in sample 4L and 5P. *Proteobacteria* was affiliated to *Alphaproteobacteria*, *Betaproteobacteria,* and *Gamaproteobacteria* (Figure 3.2A).

Differences among samples were obviously seen at family level of taxonomic classification (Figure 3.2B). Sample 1F, 2R and 3S supplied by the same source water, showed similar pattern. And sample 4L and 5P have shared more similarity with their abundant families. Sample 1F, 2R and 3S were dominant by *Caulobacteraceae* (25.0%, 25.2% and 14.5%), *Sphingomonadaceae* (17.7%, 30.4% and 6.2%), and *Oxalobacteraceae* (14.7%, 24.3% and 48.5%). Family members in *Sphingomonadaceae* (23.5% and 9.8%), *Moraxellaceae* (15.7% and 9.9%), *Methylobacteriaceae* (15.8% and 2.8%), and *Oxalobacteraceae* (7.0% and 6.4%) had high abundance in both sample 4L and 5P. High abundance of *Mycobacteriaceae* (23.6%) was found in sample 5P.

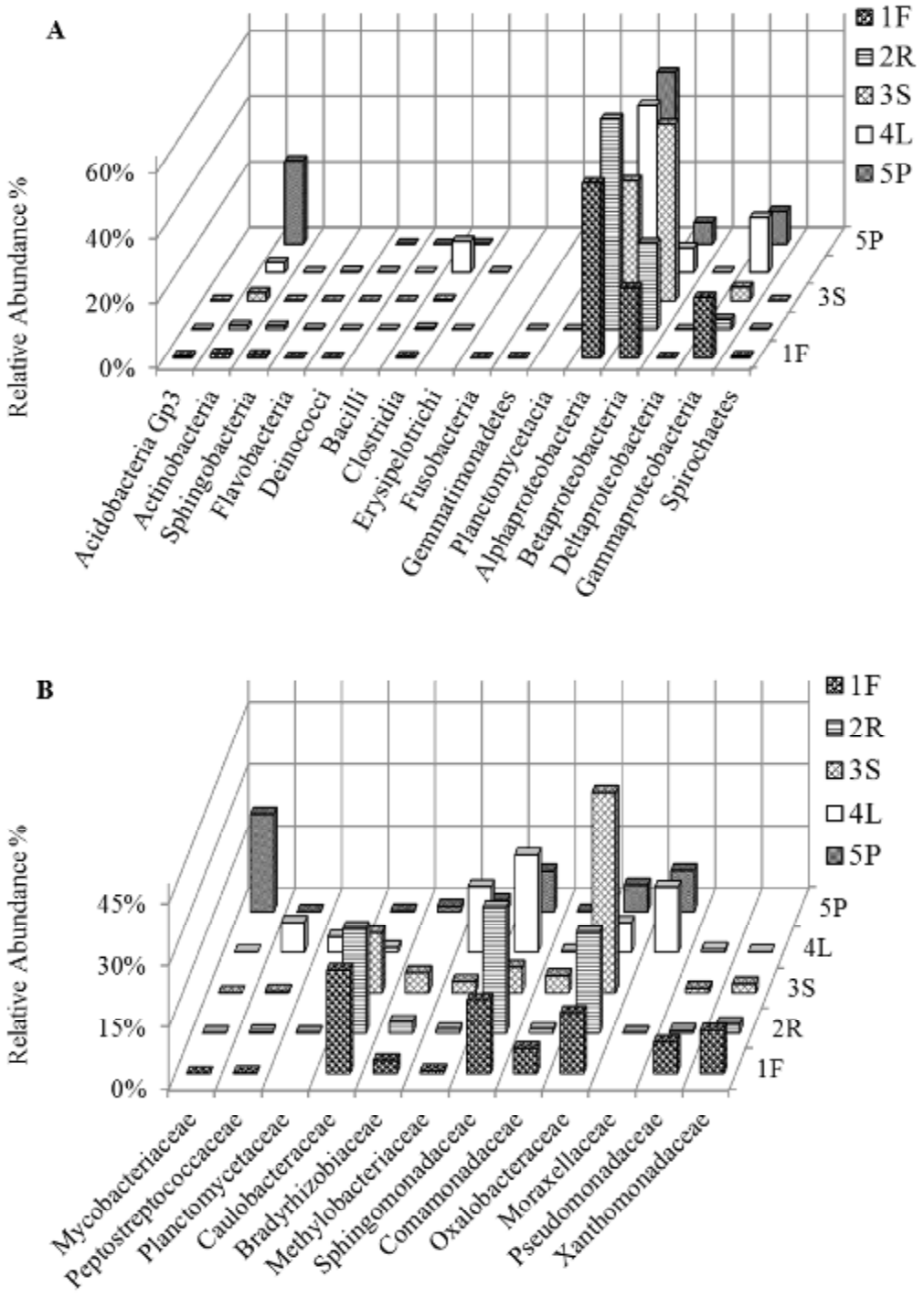

Figure 3.2 Relative abundances of bacterial classes (A) (> 1%) and dominant families (B) (> 2%) in five different drinking water samples as revealed by pyrosequencing analysis.

To verify the pyrosequencing results, parallel clone library analyses with an average length of 800 bp were also performed for water sample 2R, 3S and 5P. Total 267 clones were analyzed from three water samples, which represent 55 unique sequences. The abundant bacterial families identified by clone library analyses were consistent with the corresponding pyrosequencing analyses as shown in Figure 3.3. *Oxalobacteraceae* (31.8% and 23.5%), *Sphingomonadaceae* (16.9% and 13.5%), *Mycobacteriaceae* (12.0% and 10.8%), *Caulobacteraceae* (7.1% and 10.4%) and *Moraxellaceae* (9.4% and 4.5%) were the top 5 families identified by both clone library and pyrosequencing analyses. In clone library data set 88% of the sequences can be classified to genus level. The phylogenetic distribution of bacteria in drinking water samples was shown in Figure 3.4. Abundant genera identified by clone library analyses affiliated with the above top 5 families were *Massilia* (30.3%), *Sphingomonas* (10.9%), *Mycobacterium* (12.0%), *Caulobacter* (4.5%), *Brevundimonas* (2.6%), and *Acinetobacter* (9.4%).

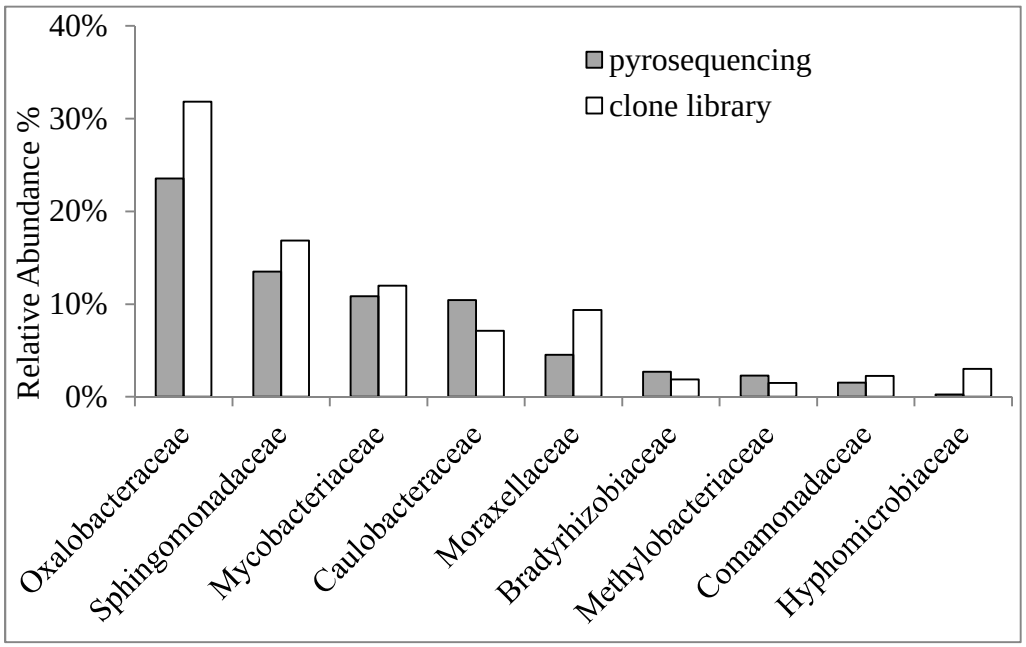

Figure 3.3 Comparison of total bacterial community compositions identified by pyrosequencing and clone library at the family level in drinking water 2R, 3S and 5P: Shown are bacterial taxa with relative abundance greater than 2%.

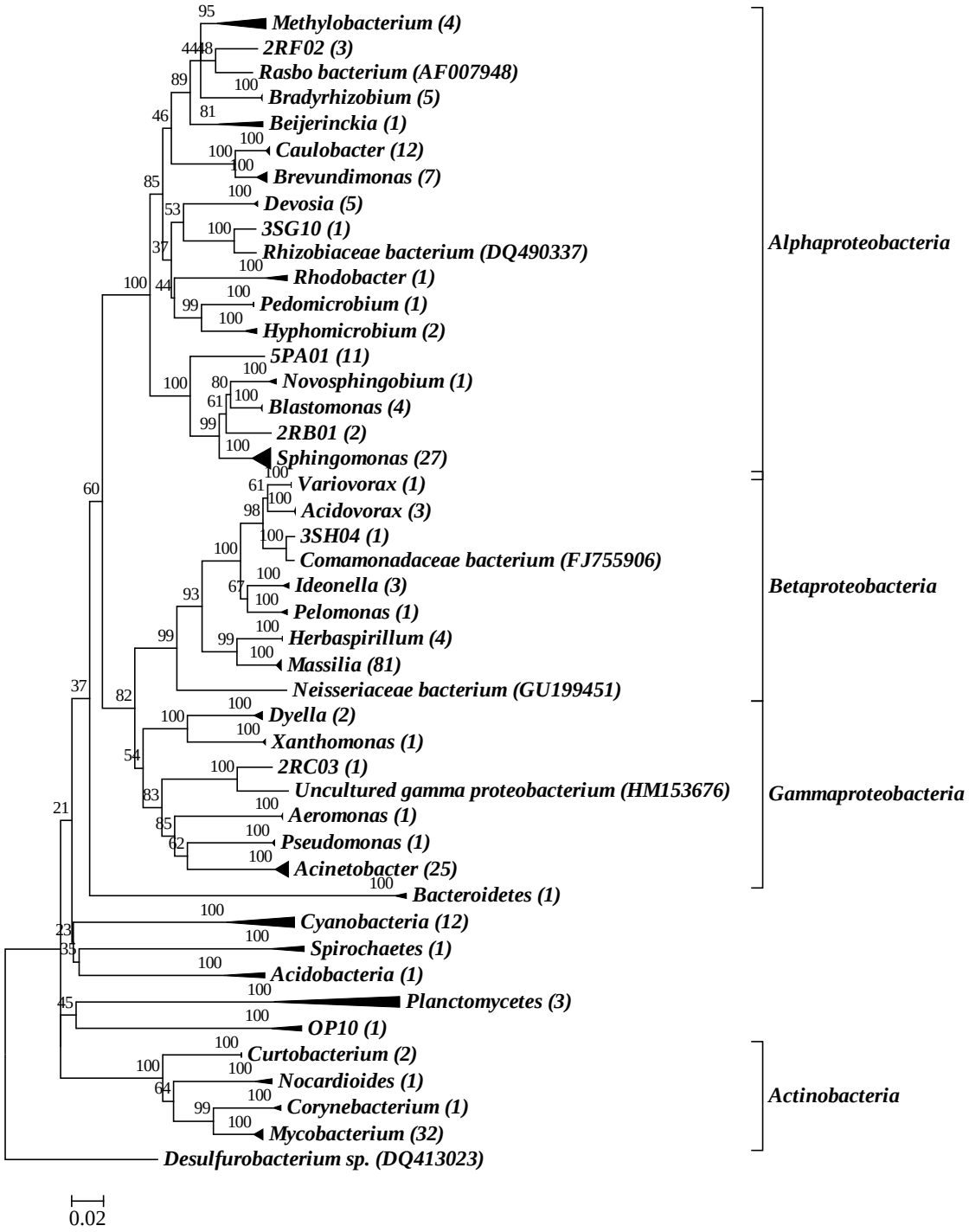

Figure 3.4 Neighbor-joining phylogenetic tree of 16S rDNA sequences from 3 drinking water clone libraries and their closest known relatives. The numbers at the nodes indicate the percentages of occurrence in 1000 bootstrapped. The numbers in the parentheses indicate the occurrence of specific OTU in sampled libraries.

The most abundant family group of sample 5P (35.1%) in pyrosequencing data set couldn't be assigned to a known family member, however identified as *Alphaproteobacteria* by RDP database. The corresponding sequences in clone library data set were fished out by comparing the V3 region of clone library *Alphaproteobacteria* clones with the unclassified *Alphaproteobacteria* sequences in pyrosequencing dataset. The corresponding representative clone (5PA01) were assigned to *Sphingomonadaceae* family, however was unclassifiable at genus level. The NCBI blast results showed that the representative clone have 92% identity with *Sphingomonas sanxanigenens* strain NX02 (DQ789172), 99% identity with uncultured bacterium clone B1NR70D5 (AY957890) which was found in drinking water biofilm (Williams et al. 2005). The phylogenetic relationships of the unclassified *Alphaproteobacteria* with their close relatives were shown in Figure 3.5.

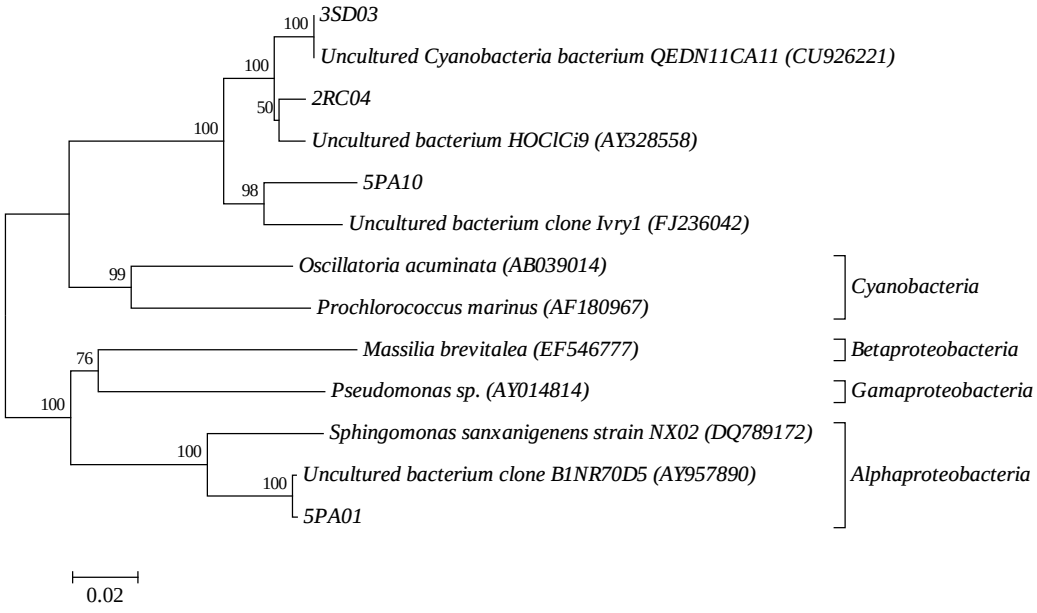

Figure 3.5 Neighbor-joining phylogenetic tree of unclassified 16S rDNA sequences from 3 drinking water clone libraries and their closest known relatives. The numbers at the nodes indicate the percentages of occurrence in 1000 bootstrapped. The numbers in the parentheses indicate the occurrence of specific OTU in sampled libraries.

Another group in clone library data set, accounted for 4.5% of the total clone library data was unable to be classified by RDP classifier, however was assigned to *Cyanobacteria* group A and group B by other researchers (Poitelon et al. 2009, Williams et al. 2005). The unclassified sequences in sample 2R and 3S showed 98% identity with HOClCi9 (AY328558) (group A) found in drinking water biofilm (Williams et al. 2004). The unclassified sequences in sample 5P had only 92% identity with Ivry1 (FJ236042) (group B) (Figure 3.5). These unclassified sequences may belong to potential novel linages with no represent cultures were described. Molecular techniques allow the detection of unknown bacteria. More work need to be done to examine their phylogenetic position and their potential function in drinking water. In summary, the bacteria taxonomic distribution of this survey was agreed with other reports for bacteria identified in drinking water (Hong et al. 2010, Revetta et al. 2011, White et al. 2011). Despite the short read, pyrosequencing generated much more sequences than clone library and other 16S rRNA based molecular analysis, which provides a great opportunity to estimate bacterial diversity for drinking water.

3.4.2. Bacterial diversity in drinking water

The bacterial diversity in drinking water was evaluated based on pyrosequencing analysis using the taxonomic classified phylotypes instead of taxonomic units based on percentage similarity. The rarefaction curves approximately reached to a plateau for all water samples at genus level, suggesting that enough bacterial genera were covered in this survey (Figure 3.6). The number of bacteria genera observed in sample 4L and 5P were less than the other three samples (Table 3.1). Sample 5P had about half of bacterial genera numbers observed in 4L and other samples. Most bacteria in sample 5P can be detected in 4L. 77.3% genera and 89.7% families accounted for

98.6% of the total population in sample 5P can be detected in 4L. The species richness estimators (Chao1 and ACE) and diversity indices also showed that water sample 5P had a lower bacterial diversity than the other four samples. Instead of using traditional sand filtration, water treatment plant for sample 5P employed microfiltration membrane to remove particles and microbes. The membrane greatly decrease bacterial diversity in drinking water compared with water treated through sand filtration.

3.4.3. Core population and pathogen signature in drinking water

Taxonomic analysis revealed a broad range of bacteria widely distributed in drinking water samples. Some core populations were found persistent in all water samples based on pyrosequencing analysis. *Alphaproteobacteria, Betaproteobacteria,* and *Gamaproteobacteria* were dominant and wide spread in all drinking water samples. Previous study on fresh water and drinking water bacteria also showed the dominant of *Proteobacteria* (Poitelon et al. 2009, Zwart et al. 2002). *Alphaproteobacteria* and *Betaproteobacteria* have also been found to be the primary constituents of biofilm in some water distribution systems (Hong et al. 2010, Mathieu et al. 2009, Yu et al. 2010), presenting the possibility that biofilm might also serve as a source of microorganisms to drinking water. Relative low abundance of *Actinobacteria, Firmicutes, Bacteroidetes,* and *Deinococcus-Thermus* were also observed across all samples as described by other researches (Hong et al. 2010, Revetta et al. 2011).

Table 3.1 Summary of sequencing reads from pyrosequencing of 16S rDNA gene amplicons

Sample	No. of sequences	Observed genera	Shannon[a]	Chao1[a]	ACE[a]
1F	4996	88	2.79 (2.75 - 2.83)	101 (92 - 132)	100 (93 - 119)
2R	5136	83	2.43 (2.39 - 2.47)	109 (92 - 160)	103 (91 - 130)
3S	6793	82	2.29 (2.24 - 2.33)	96 (86 - 125)	100 (89 - 128)
4L	5422	79	2.72 (2.68 - 2.75)	98 (86 - 134)	98 (87 - 124)
5P	10098	44	2.00 (1.97 - 2.02)	62 (49 - 104)	82 (64 - 119)

[a]Numbers in parentheses indicate the lower and upper bounds of 95% confidence interval.

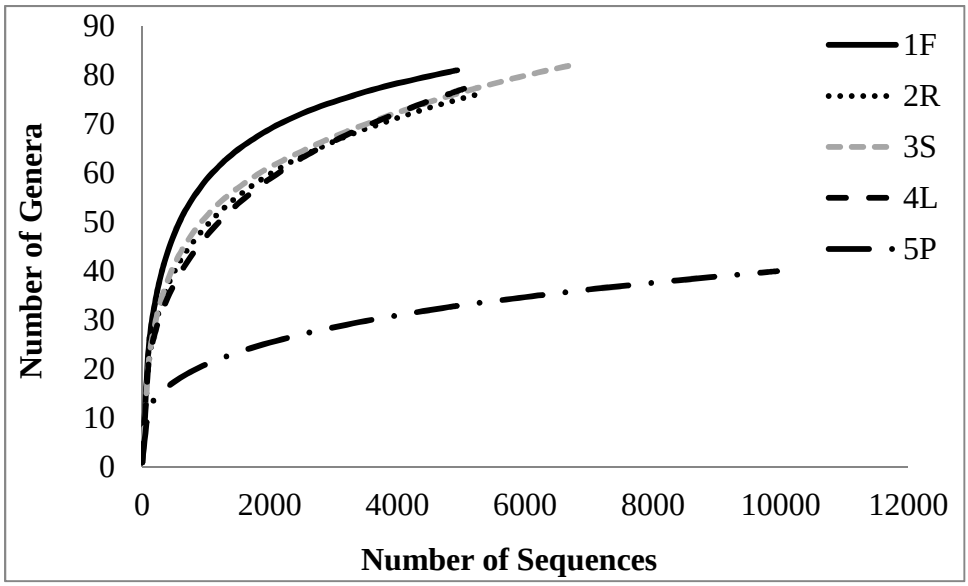

Figure 3.6 Rare faction curves of bacterial communities from five different drinking water samples assessed with pyrosequencing analysis.

Total 20 bacteria families shared by all the water samples were observed in this survey. Seven abundant core families (above 2% abundance in at least one water sample) spread in all drinking water samples: *Sphingomonadaceae, Caulobacteraceae, Methylobacteriaceae, Oxalobacteraceae, Comamonadaceae, Mycobacteriaceae,* and *Peptostreptococcaceae*. These families represented by *Sphingomonas*, *Novosphingobium*, *Caulobacter*, *Methylobacterium, Massilia*, *Acidovorax,* and *Mycobacterium*. *Massilia* was found dominant in lead corroded drinking water pipe biofilm (White et al. 2011). *Massilia* are capable of using a variety of organic compounds including volatile fatty acids and exhibited the tendency to form pellicles in liquid medium, suggesting the potential of biofilm formation (Gallego et al. 2006b, Zu et al. 2008). *Sphingomonas* and *Acidovorax* were reported as dominant population in water meter biofilm (Hong et al. 2010). Bacteria from these two genera are capable of using a wide range of substrates and therefore ubiquitous in various environments (Balkwill et al. 1997, Takeuchi et al.

1995, Willems et al. 1990). *Novosphingobium* and *Caulobacter* were observed in the occurrence of red water (Li et al. 2010). *Novosphingobium* a member of *Sphingomonadaceae* has many common features with *Sphingomonas*. *Caulobacter* is well known to be capable of surface attachment and biofilm formation (Entcheva-Dimitrov and Spormann 2004). Bacteria from *Mycobacterium* are notorious waterborne pathogens and some were found tolerant to chlorine (Le Dantec et al. 2002, Whipps et al. 2007). The unique cell wall and their ability to form biofilm make them persistent in drinking water (Torvinen et al. 2007). *Methylobacterium* was also a common resident of drinking water (Berg et al. 2009). *Methylobacterium* isolates from drinking water are capable of forming biofilm and utilizing a diverse group of carbon substrates in addition to C1 compounds (Gallego et al. 2005, 2006a, Simoes et al. 2010). These common features, such as versatility in substrate utilization and potential in biofilm formation, shared by the core populations in drinking water may allow them persistent in a harsh environment.

Potential pathogens were found in the genera of *Mycobacterium*, *Staphylococcus*, *Clostridium*, *Escherichia/Shigella*, *Aeromonas*, *Legionella*, *Stenotrophomonas*, *Leptospira,* and *Sporacetigenium*. *Mycobacterium* was found in all the water samples accounted for 23.6% of sample 5P and less than 0.06% in all the other samples. Other pathogens were not observed in all the samples and showed very low abundance (less than 0.1%). *Mycobacterium* was the only pathogen found in sample 5P. However this survey was based on DNA analysis the dead cell may also be detected and the activities and the potential risk of those pathogens are not very clear.

3.4.4. Link between bacterial communities and environment

To explain the compositional variations in drinking water, water quality were measured and principal component analysis were perform to show the link between bacterial communities and environmental variables. The water quality data were similar to the corresponding water quality report data posted by water treatment plant and in the regulation range maintained in drinking water distribution systems in US. As shown in Table 3.2, free chlorine ranged from 3.13 mg/l to 0.71 mg/l. Less numbers of heterotrophic bacteria were observed in water samples with high free chlorine than those with lower free chlorine concentration. The HPC number in this survey was far below the USEPA regulation rules for drinking water (500 CFU/ml) and lower than most reported drinking water survey with a range of 2 to 10^4 CFU/ml (Kahlisch et al. 2011, Lautenschlager et al. 2010, Pepper et al. 2004).

Table 3.2 Water quality parameters

Sample	1F	2R	3S	4L	5P
pH	7.33 ± 0.01	7.49 ± 0.02	6.77 ± 0.02	7.32 ± 0.02	7.17 ± 0.02
Turbidity	0.078 ± 0.000	0.024 ± 0.000	0.010 ± 0.002	0.039 ± 0.00	0.033 ± 0.000
DOC (mg/L)	1.64 ± 0.02	1.14 ± 0.03	1.26 ± 0.02	2.99 ± 0.04	2.25 ± 0.03
Free Cl_2 (mg/L)	0.75 ± 0.01	0.71 ± 0.06	2.13 ± 0.03	3.13 ± 0.08	1.55 ± 0.01
Hardness (mg/L as $CaCO_3$)	103 ± 4	155 ± 5	67 ± 2	145 ± 4	180 ± 2
Conductivity (uS/cm)	258 ± 1	358 ± 1	260 ± 1	336 ± 0	377 ± 1
chloride (mg/L)	13.22 ± 0.07	17.20 ± 0.28	10.93 ± 0.56	13.09 ± 0.03	13.29 ± 0.33
Sulfate (mg/L)	7.60 ± 0.22	10.85 ± 0.42	16.93 ± 0.01	23.97 ± 0.37	8.77 ± 0.09
Nitrate (mg/L)	0.54 ± 0.04	3.21 ± 0.07	1.29 ± 0.01	2.54 ± 0.03	0.91 ± 0.00
HPC (CFU/ml)	28.0 ± 2.6	23.7 ± 1.4	2.9 ± 0.2	4.96 ± 0.58	18.9 ± 1.8

Differences of bacterial community composition in water samples were observed at family level analysis (Figure 3.2B). Therefore PCA analysis was performed with selected environmental variables and bacterial family abundance. PCA diagram suggested that most of bacterial families contribute to PC1, *Mycobacteriaceae* contribute mostly to PC2 (Figure 3.7). PC1 represent 51.4% bacterial composition variables and total 76.8% of the variables can be explained by the first two components. Sample 4L and 5P were separate from the other three samples along PC1. Sample 5P was separated from other samples along PC2 because the dominant family *Mycobacteriaceae* in 5P had relatively low abundance in other samples. Sample 1F, 2R, and 3S were clustered together because they shared relatively high abundance families appeared in the left corner of PCA diagram. Free chlorine and DOC vectors have very small angle with PC1 axis suggested strong correlation between the two environmental variables and bacterial communities represented by PC1 ($r = 0.73$ and 0.89 for free chlorine and DOC, respectively). The PCA result was confirmed by cluster analysis of bacterial community. Hierarchical cluster analysis of pyrosequencing fingerprints based on weighted Unifrac distance also showed that bacterial communities in drinking water were influenced by source water as great differences were observed for drinking water samples treated from different river water (Figure 3.1 and 8). Despite the differences of sample 5P and 4L in their community structure pattern as revealed by the relative abundance of different populations, 5P and 4L are different sample 5P and 4L shared a large amount of common phylotypes (77.3% genera and 89.7% families accounted for 98.6% of the total population in sample 5P). Therefore sample 5P and 4L were grouped together when the tree branches were weighted by the amount of population.

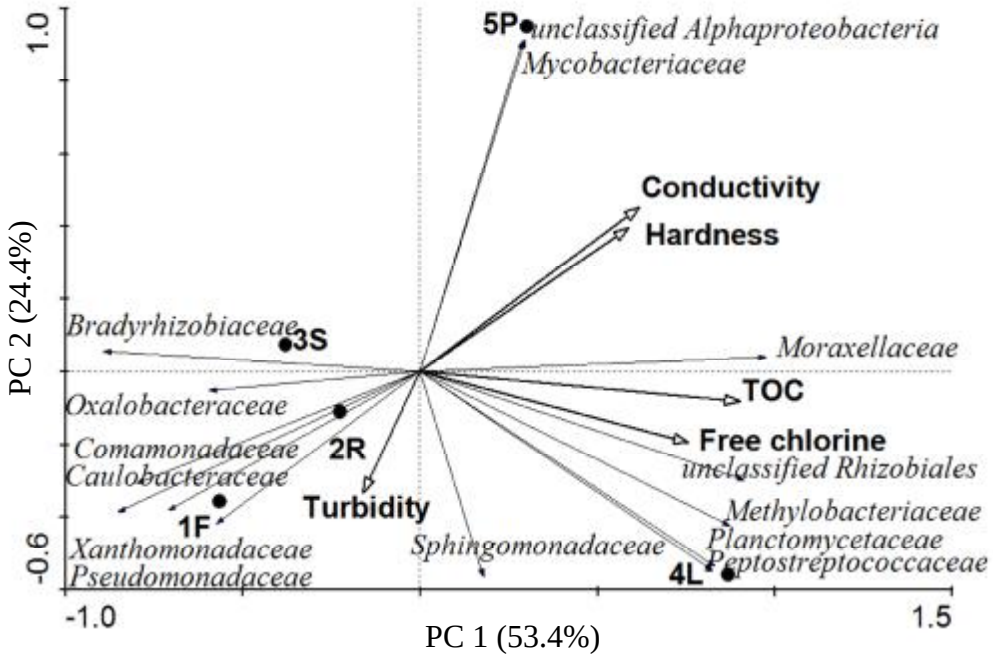

Figure 3.7 Principal component analysis (PCA) of bacterial communities and drinking water quality parameters based on bacterial families with relative abundance above 2%. Every vectors point to the direction of increase for a given variable so that water samples with similar communities are located in the similar positions in the diagram.

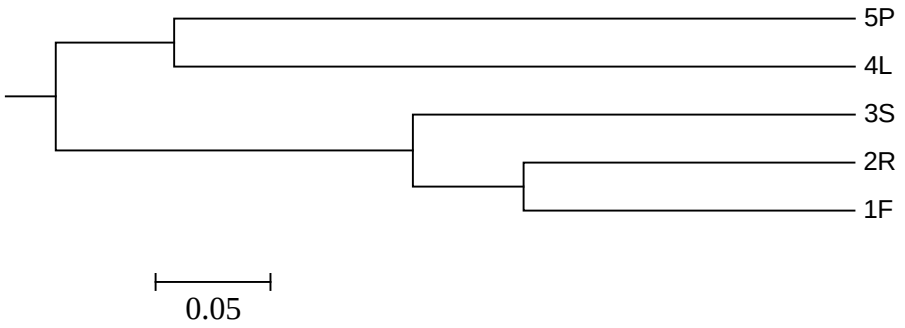

Figure 3.8 Hierarchical clustering of bacterial community from five drinking water samples assessed with pyrosequencing analysis. The bar represents a weighted UniFrac distance of 0.05

Previous studies demonstrated that bacterial communities in tap water can be affected by many factors such as source water, treatment process, drinking water quality and the formation of biofilm in water distribution pipes (Eichler et al. 2006, Poitelon et al. 2010, Stewart et al. 1990). Tap water bacterial community compositions were found similar to their source reservoirs by other researchers (Eichler et al. 2006, Poitelon et al. 2010). Higher level of DOC and lack of chlorine residual may lead to the changes of different *Proteobacteria* population (Kalmbach et al. 1997, Lechevallier et al. 1987, Williams et al. 2004). *Betaproteobacteria* was considered more sensitive to chlorine than *Alphaproteobacteia* (Kormas et al. 2010, Williams et al. 2004). In our survey less number of phylotyes was identified from *betaproteobacteia* (10 families) than from *Alphaproteobacteia* (19 families) in the water samples. And most samples have higher relative abundance of *Alphaproteobacteia* than *Betaproteobacteia* (Figure 3.2A). However, *Betaproteobacteia* was also observed as predominant population in biofilm samples (Kalmbach et al. 1997, Schwartz et al. 2003). The presence of biofilm may protect many bacteria survive the disinfection exposure. Therefore, the bacterial community pattern in drinking water affected by the interactions of many environmental variables.

3.5. Conclusions

Bacterial community compositions in drinking water were influenced by source water and environmental variables. Drinking water harbors a diverse bacterial community as revealed by high-throughput pyrosequencing. Core populations from *Alphaproteobacteria*, *Betaproteobacteria,* and *Actinobacteria* were identified in all the samples. Core abundant families shared by all the water samples were *Sphingomonadaceae*, *Caulobacteraceae*, *Methylobacteriaceae*, *Oxalobacteraceae*, *Comamonadaceae*, *Mycobacteriaceae* and *Peptostreptococcaceae*, which represented by *Sphingomonas*, *Novosphingobium*, *Caulobacter*,

Methylobacterium, *Massilia*, *Acidovorax*, and *Mycobacterium*.

3.6. References

APHA (2005) Standard methods for the examination of water and wastewater., American Public Health Association, Washington, D.C.

Balkwill, D.L., Drake, G.R., Reeves, R.H., Fredrickson, J.K., White, D.C., Ringelberg, D.B., Chandler, D.P., Romine, M.F., Kennedy, D.W. and Spadoni, C.M. (1997) Taxonomic study of aromatic-degrading bacteria from deep-terrestrial-subsurface sediments and description of Sphingomonas aromaticivorans sp nov, Sphingomonas subterranea sp nov, and Sphingomonas stygia sp nov. International Journal of Systematic Bacteriology 47(1), 191-201.

Berg, K.A., Lyra, C., Sivonen, K., Paulin, L., Suomalainen, S., Tuomi, P. and Rapala, J. (2009) High diversity of cultivable heterotrophic bacteria in association with cyanobacterial water blooms. Isme Journal 3(3), 314-325.

Costello, E.K., Lauber, C.L., Hamady, M., Fierer, N., Gordon, J.I. and Knight, R. (2009) Bacterial Community Variation in Human Body Habitats Across Space and Time. Science 326(5960), 1694-1697.

Craun, G.F., Brunkard, J.M., Yoder, J.S., Roberts, V.A., Carpenter, J., Wade, T., Calderon, R.L., Roberts, J.M., Beach, M.J. and Roy, S.L. (2010) Causes of Outbreaks Associated with Drinking Water in the United States from 1971 to 2006. Clinical Microbiology Reviews 23(3), 507-528.

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.

Entcheva-Dimitrov, P. and Spormann, A.M. (2004) Dynamics and control of biofilms of the oligotrophic bacterium Caulobacter crescentus. Journal of Bacteriology 186(24), 8254-8266.

Gallego, V., Garcia, M.T. and Ventosa, A. (2005) Methylobacterium hispanicum sp nov and Methylobacterium aquaticum sp nov., isolated from drinking water. International Journal of Systematic and Evolutionary Microbiology 55, 281-287.

Gallego, V., Garcia, M.T. and Ventosa, A. (2006a) Methylobacterium adhaesivum sp nov, a methylotrophic bacterium isolated from drinking water. International Journal of Systematic and Evolutionary Microbiology 56, 339-342.

Gallego, V., Sanchez-Porro, C., Garcia, M.T. and Ventosa, A. (2006b) Massilia aurea sp nov., isolated from drinking water. International Journal of Systematic and Evolutionary Microbiology 56, 2449-2453.

Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J. and Knight, R. (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nature Methods 5(3), 235-237.

Hill, V.R., Kahler, A.M., Jothikumar, N., Johnson, T.B., Hahn, D. and Cromeans, T.L. (2007) Multistate evaluation of an ultrafiltration-based procedure for simultaneous recovery of enteric microbes in 100-liter tap water samples (vol 73, pg 4218, 2007). Applied and Environmental Microbiology 73(19), 6327-6327.

Hong, P.Y., Hwang, C.C., Ling, F.Q., Andersen, G.L., LeChevallier, M.W. and Liu, W.T. (2010) Pyrosequencing Analysis of Bacterial Biofilm Communities in Water Meters of a Drinking Water Distribution System. Applied and Environmental Microbiology 76(16), 5631-5635.

Huse, S.M., Dethlefsen, L., Huber, J.A., Welch, D.M., Relman, D.A. and Sogin, M.L. (2008) Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. Plos Genetics 4(11).

Huse, S.M., Welch, D.M., Morrison, H.G. and Sogin, M.L. (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environmental Microbiology 12(7), 1889-1898.

Jjemba, P.K., Weinrich, L.A., Cheng, W., Giraldo, E. and LeChevallier, M.W. (2010) Regrowth of Potential Opportunistic Pathogens and Algae in Reclaimed-Water Distribution Systems. Appl. Environ. Microbiol. 76(13), 4169-4178.

Kahlisch, L., Henne, K., Gröbe, L., Brettar, I. and Höfle, M. (2011) Assessing the Viability of Bacterial Species in Drinking Water by Combined Cellular and Molecular Analyses. Microbial Ecology, 1-15.

Kalmbach, S., Manz, W. and Szewzyk, U. (1997) Dynamics of biofilm formation in drinking water: Phylogenetic affiliation and metabolic potential of single cells assessed by formazan reduction and in situ hybridization. Fems Microbiology Ecology 22(4), 265-279.

Kormas, K.A., Neofitou, C., Pachiadaki, M. and Koufostathi, E. (2010) Changes of the bacterial assemblages throughout an urban drinking water distribution system. Environmental Monitoring and Assessment 165(1-4), 27-38.

Lane, D.J. (1990) 16S/23S rRNA sequencing, p. 115-147. *In* E. Stackebrandt and M. Goodfellow (ed.), Nucleic acid techniques in bacterial systematics., John Wiley & Sons, Chichester, United Kingdom.

Lautenschlager, K., Boon, N., Wang, Y.Y., Egli, T. and Hammes, F. (2010) Overnight stagnation of drinking water in household taps induces microbial growth and changes in community composition. Water Research 44(17), 4868-4877.

Le Dantec, C., Duguet, J.P., Montiel, A., Dumoutier, N., Dubrou, S. and Vincent, V. (2002) Occurrence of mycobacteria in water treatment lines and in water distribution systems. Applied and Environmental Microbiology 68(11), 5318-5325.

Leach, L.H., Zhang, P., LaPara, T.M., Hozalski, R.M. and Camper, A.K. (2009) Detection and enumeration of haloacetic acid-degrading bacteria in drinking water distribution systems using dehalogenase genes. Journal of Applied Microbiology 107(3), 978-988.

Lechevallier, M.W., Babcock, T.M. and Lee, R.G. (1987) Examination and characterization of distribution system biofilms. Applied and Environmental Microbiology 53(12), 2714-2724.

Li, D., Li, Z., Yu, J., Cao, N., Liu, R. and Yang, M. (2010) Characterization of bacterial community structure in a drinking water distribution system during an occurrence of red water. Appl. Environ. Microbiol. 76(21), 7171-7180.

Lipponen, M.T.T., Martikainen, P.J., Vasara, R.E., Servomaa, K., Zacheus, O. and Kontro, M.H. (2004) Occurrence of nitrifiers and diversity of ammonia-oxidizing bacteria in developing drinking water biofilms. Water Research 38(20), 4424-4434.

Martin-Laurent, F., Philippot, L., Hallet, S., Chaussod, R., Germon, J.C., Soulas, G. and Catroux, G. (2001) DNA extraction from soils: Old bias for new microbial diversity analysis methods. Applied and Environmental Microbiology 67(5), 2354-2359.

Martiny, A.C., Albrechtsen, H.J., Arvin, E. and Molin, S. (2005) Identification of bacteria in biofilm and bulk water samples from a nonchlorinated model drinking water distribution system: Detection of a large nitrite-oxidizing population associated with Nitrospira spp. Applied and Environmental Microbiology 71(12), 8611-8617.

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

Mull, B. and Hill, V.R. (2009) Recovery and Detection of Escherichia coli O157:H7 in Surface Water, Using Ultrafiltration and Real-Time PCR. Applied and Environmental Microbiology 75(11), 3593-3597.

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.

Pepper, I.L., Rusin, P., Quintanar, D.R., Haney, C., Josephson, K.L. and Gerba, C.P. (2004) Tracking the concentration of heterotrophic plate count bacteria from the source to the consumer's tap. International Journal of Food Microbiology 92(3), 289-295.

Poitelon, J.B., Joyeux, M., Welte, B., Duguet, J.P., Prestel, E., Lespinet, O. and Dubow, M.S. (2009) Assessment of phylogenetic diversity of bacterial microflora in drinking water using serial analysis of ribosomal sequence tags. Water Research 43(17), 4197-4206.

Poitelon, J.B., Joyeux, M., Welte, 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 & Biotechnology 37(2), 117-128.

Polaczyk, A.L., Narayanan, J., Cromeans, T.L., Hahn, D., Roberts, J.M., Amburgey, J.E. and Hill, V.R. (2008) Ultrafiltration-based techniques for rapid and simultaneous concentration of multiple microbe classes from 100-L tap water samples. Journal of Microbiological Methods 73(2), 92-99.

Reasoner, D.J. and Geldreich, E.E. (1985) A new medium for the enumeration and subculture of bacteria from potable water. Applied and Environmental Microbiology 49(1), 1-7.

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.

Revetta, R.P., Matlib, R.S. and Domingo, J.W.S. (2011) 16S rRNA Gene Sequence Analysis of Drinking Water Using RNA and DNA Extracts as Targets for Clone Library Development. Current Microbiology 63(1), 50-59.

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., Stres, B., Thallinger, G.G., Van Horn, D.J. and Weber, C.F. (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology 75(23), 7537-7541.

Schwartz, T., Hoffmann, S. and Obst, U. (2003) Formation of natural biofilms during chlorine dioxide and u.v. disinfection in a public drinking water distribution system. Journal of Applied Microbiology 95(3), 591-601.

Simoes, L.C., Simoes, M. and Vieira, M.J. (2010) Adhesion and biofilm formation on polystyrene by drinking water-isolated bacteria. Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology 98(3), 317-329.

Stewart, M.H., Wolfe, R.L. and Means, E.G. (1990) Assessment of the bacteriological activity associated with granular activated carbon treatment of drinking water. Applied and Environmental Microbiology 56(12), 3822-3829.

Takeuchi, M., Sakane, T., Yanagi, M., Yamasato, K., Hamana, K. and Yokota, A. (1995) Taxonomic Study of Bacteria Isolated from Plants: Proposal of Sphingomonas rosa sp. nov., Sphingomonas pruni sp. nov., Sphingomonas asaccharolytica sp. nov., and Sphingomonas mali sp. nov. . International Journal of Systematic Bacteriology 45(2), 334-341.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution 28(10), 2731-2739.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25(24), 4876-4882.

Torvinen, E., Lehtola, M.J., Martikainen, P.J. and Miettinen, I.T. (2007) Survival of Mycobacterium avium in drinking water biofilms as affected by water flow velocity, availability of phosphorus, and temperature. Applied and Environmental Microbiology 73(19), 6201-6207.

Wang, Q., Garrity, G.M., Tiedje, J.M. and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73(16), 5261-5267.

Whipps, C.M., Butler, W.R., Pourahmad, F., Watral, V.G. and Kent, M.L. (2007) Molecular systematics support the revival of Mycobacterium salmoniphilum (ex Ross 1960) sp nov., nom. rev., a species closely related to Mycobacterium chelonae. International Journal of Systematic and Evolutionary Microbiology 57, 2525-2531.

White, C., Tancos, M. and Lytle, D.A. (2011) Microbial Community Profile of a Lead Service Line Removed from a Drinking Water Distribution System. Applied and Environmental Microbiology 77(15), 5557-5561.

Will, C., Thurmer, A., Wollherr, A., Nacke, H., Herold, N., Schrumpf, M., Gutknecht, J., Wubet, T., Buscot, F. and Daniel, R. (2010) Horizon-Specific Bacterial Community Composition of German Grassland Soils, as Revealed by Pyrosequencing-Based Analysis of 16S rRNA Genes. Appl. Environ. Microbiol. 76(20), 6751-6759.

Willems, A., Falsen, E., Pot, B., Jantzen, E., Hoste, B., Vandamme, P., Gillis, M., Kersters, K. and Deley, J. (1990) Acidovorax, a new genus for Pseudomonas facilis, Pseudomonas delafieldii, E. Falsen (EF) group 13, EF group 16, and several clinical isolates, with the species Acidovorax facilis comb. nov., Acidovorax delafieldii comb. nov., and Acidovorax temperans sp. nov. International Journal of Systematic Bacteriology 40(4), 384-398.

Williams, M.M., Domingo, J.W.S., Meckes, M.C., Kelty, C.A. and Rochon, H.S. (2004) Phylogenetic diversity of drinking water bacteria in a distribution system simulator. Journal of Applied Microbiology 96(5), 954-964.

Williams, M.M., Domingo, J.W.S. and Meckes, M.C. (2005) Population diversity in model potable water biofilms receiving chlorine or chloramine residual. Biofouling 21(5-6), 279-288.

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.

Zhang, P., Hozalski, R.M., Leach, L.H., Camper, A.K., Goslan, E.H., Parsons, S.A., Xie, Y.F.F. and LaPara, T.M. (2009) Isolation and characterization of haloacetic acid-degrading Afipia spp. from drinking water. Fems Microbiology Letters 297(2), 203-208.

Zhang, Y., Zamudio Canas, E.M., Zhu, Z.W., Linville, J.L., Chen, S. and He, Q. (2011) Robustness of archaeal populations in anaerobic co-digestion of dairy and poultry wastes. Bioresource Technology 102(2), 779-785.

Zu, D., Wanner, G. and Overmann, J. (2008) Massilia brevitalea sp nov., a novel betaproteobacterium isolated from lysimeter soil. International Journal of Systematic and Evolutionary Microbiology 58, 1245-1251.

Zwart, G., Crump, B.C., Agterveld, M., Hagen, F. and Han, S.K. (2002) Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. Aquatic Microbial Ecology 28(2), 141-155.

Chapter 4. Bacterial Community Dynamics during Drinking Water Treatment and Distribution Processes

4.1. Abstract

Bacterial community dynamics during drinking water treatment and distribution processes were investigated by pyrosequencing of the 16S rRNA amplicons at two different surface water treatment plants: a membrane filtration plant and a conventional sand filtration plant. Water samples were taken from source river water to customer's tap water after different steps of treatment (filtration, chlorination, distribution, and stagnation). Substantial differences were observed after each treatment and distribution steps for both treatment plants. Membrane filtration removed a large variety of bacteria populations; however, was less effective in removing bacteria from the genus of *Delftia* and *Pseudomonas*. Chlorine disinfection was the key step for bacteria removal, subsequently playing an important role in shaping bacteria community structure in tap water. Conventional sand filtration and disinfection also greatly affected the bacterial community composition in water. Variety of bacteria passed though the unspecific sieving of sand filters. The distribution had greater influence on the fresh tap water than disinfection. Bacteria from genus of *Pseudomonas*, *Stenotrophomonas*, *Methylobacterium*, *Massilia*, *Naxibacter*, *Undibacterium*, and *Acidovorax* were dominated in finished water sample suggesting that these bacteria genera might be potential chlorine resistant population that can survive disinfection step. Stagnation showed substantial influence on the bacterial community composition for both water treatment plants. After stagnation, *Betaproteobacteria* greatly decreased, instead *Alphaproteobacteria* and *Firmicutes* greatly increased and became the dominant population. During the whole process for drinking water purification and distribution, bacteria members from the family of *Sphingomonadaceae*, *Burkholderiaceae*, *Comamonadaceae*, *Oxalobacteraceae*, and *Xanthomonadaceae* were detected with high frequency in multiple

treatment steps in both water treatment plants. Bacteria members from these families were the core resilient populations which have high chance to survive treatment and distribution process, therefore worth of developing further control strategy. Bacteria members from the genus of *Methylobacterium*, *Sphingomonas*, *Paracoccus,* and *Mycobacterium* were isolated in the tap water and stagnated water. The occurrence of these bacteria at customer's tap indicated potential risk for human health cause by distribution and stagnation.

4.2. Introduction

The ultimate goal of drinking water treatment is to remove the physical, chemical and biological contaminants in the source water. Conventional treatment use sand filters to remove coagulated organic matters, particles including some microorganisms in the source water. Then disinfectant were added to the water and maintained in certain level to kill microorganisms or inhibit their proliferation. However, no matter what kinds of filter media or disinfectants were used, there are always bacteria capable of escaping the filter screens and surviving the disinfection process (Eichler et al. 2006, Ho et al. 2012, Poitelon et al. 2010). The regrowth of the bacterial may lead to serious health risk for human (Craun et al. 2010).

In recent years a new membrane filtration technology was employed as an attractive alternative to conventional sand filtration due to their good reputation for turbidity and microorganism removal (Ho et al. 2012). However previous studies on the membrane treatment efficiency were either performed on pilot scale or focused on the reduced bacterial number (Ho et al. 2012, Kwon et al. 2011). Few of studies did further analysis to track and identify the microorganisms that pass through the membrane sieves. In fact to our knowledge no evaluations

on the performance of membrane filtration were performed based on plant scale and whole microbial community level.

Besides, the distribution and stagnation also had considerable influence on the tap water quality (Kormas et al. 2010, Lautenschlager et al. 2010, Pepper et al. 2004). Many environmental factors changed as water pass through distribution networks and reach to the house hold pipes at customer's tap. The change of these environmental factors such as pipe material, ambient temperature, nutrient concentration and disinfection concentration had substantial influence on the regrowth of microorganisms and subsequent biofilm formation (Li et al. 2010, Pepper et al. 2004, Zhang et al. 2008). The biofilm in drinking water distribution networks was considered an important source pool for bulk water discharged at customer's tap (Henne et al. 2012, Schmeisser et al. 2003). Stagnation was reported promote the regrowth of bacteria due to the longer residence time, depletion of distinction concentration, change of temperature and pipe diameter. However most of the related studies were limited to the culturable bacteria reported as increased HPC and observation of bacteria isolates, which represented a small part of the whole bacteria community (Lautenschlager et al. 2010, Pepper et al. 2004). The dynamic changes of the whole bacterial communities during water treatment and distribution processes have not been described systematically in detail, especially for the newly applied membrane filtration process.

The aim of study is to find the key step shaping bacterial community structure in customer's drinking water, and subsequently find the core bacteria population that were able to survive treatment processes and be persistent in drinking water distribution networks. For this purpose the dynamic changes of bacterial community composition from the surface river water to the drinking water at customer's end were monitored through pyrosequencing of 16S rDNA amplicon library generated from each water samples. Two water treatment plants in local area

were selected to represent two typical treatment cases used by local drinking water supply utilities. The analysis was focused on the identification of dominant populations existed after each specific treatment distribution step. Using this approach we assessed the influence of each treatment steps on bacteria populations and zoomed in our target on several core populations occurred with high frequency in water treatment and distribution systems.

4.3. Material and Methods

4.3.1. Water sampling in two drinking water treatment plants

The two water treatment plants located in Tennessee were selected for water sampling. The sampling time for WTPA and WTPB were September and October 2012, respectively. WTPA is operated with a treatment capacity of 16 million gallons per day (MGD), on average about 8 - 10 MGD, which serves more than 24,000 people. WTPB provides finished drinking water to more than 76,000 people with a treatment capacity of over 61 million gallons per day and on an average day treats about 34 million gallons. The two plants take their surface water from different rivers, and treat through coagulation, flocculation, filtration, and chlorination. WTPA selects ZeeWeed membrane system for the filtration process, while WTPB stays with the conventional multiple-layer sand filters. Water samples taken from source to customer's tap along treatment and distribution processes were raw water (AR and BR), after filtration (AM and BS), after chlorination (AF and BF), tap water after distribution (Afr and Bfr) and overnight stagnant tap water (Ast and Bst). The tap water after distribution was taken from the same tap faucet by flushing the faucet for 10 minutes at maximum flow. The stagnant tap water was taken from the tap faucet haven't been used for 1 day (about 24 hours). Two liters of raw water and

100 liters of water after each treatment steps and tap water were collected in autoclaved carboys for biological analysis. One liter of water from each sample was used for water quality analysis.

4.3.2. Water quality analysis

Water quality parameters such as pH, turbidity, conductivity, free chlorine, dissolved organic carbon (DOC), sulfate, nitrate, and chloride were measured. Turbidity and conductivity were measured with Hach 2100N turbidimeter (Hach Company, Loveland, Colorado, USA) and Orion model 122 conductivity meter (Orion Research Inc., Boston, Massachusetts, USA), respectively. Free chlorine was quantified using "4500-Cl F" DPD Ferrous Titrimetric Method as described in standard method (APHA 2005). Dissolved organic carbon was analyzed with a Shimadzu SSM-5000A TOC analyzer (Shimadzu Corporation, Kyoto, Japan). Ions such as sulfate, nitrate, and chloride were quantified with a Dionex Ion Chromatograph ICS-2500 system (Dionex Corp., Sunnyvale, California, USA). Heterotrophic plate counts (HPC) were analyzed by incubating on R2A agar at 28°C as previously described (Reasoner and Geldreich 1985).

4.3.3. Bacteria collection and DNA extraction

Within less than 4 hours of sampling bacteria were harvested by filtration of 2 liters of raw water on a 0.2 µm pore size polycarbonate filters and 100 liters of treated water with a tangential-flow ultrafiltration system configured, prepared, and operated as previously described. All the tubes and containers included in the ultrafiltration system were disinfected with 10% hypochlorous acid, washed with deionized water, and sterilized by autoclaving before use. Chlorinated water samples were dechlorinated by the addition of sodium thiosulfate to a final concentration of 50 mg/L and sodium polyphosphate was added to each water sample to a final concentration of 0.01% (w/v) as the dispersant as previously described (Hill et al. 2007, Mull and Hill 2009,

Polaczyk et al. 2008). Bacteria were further harvested on a 0.22 μm pore size polycarbonate filters by filtration of the ultrafiltrated water. The filter sandwiches were stored at -80 °C for further analysis. The whole genome DNA was extracted from the filter sandwiches using FastDNA spin kit for soil (MP Biomedicals, Santa Anna, CA).

4.3.4. Pyrosequencing of 16S rDNA amplicons

In order to investigate the bacterial community change along the treatment process, pyrosequencing of 16S rRNA gene amplicon libraries was carried out for bacterial samples collected after different treatment steps. 16S rRNA gene amplicon libraries were generated with primers targeting the V4 hypervariable region: 515F (5'-GTGCCAGCAGCCGCGGTAA-3') and 806R (5'-GGACTACCAGGGTATCTAAT-3') (Nikkari et al. 2002). The 515F primer included a Roche 454-A pyrosequencing adapter (454 Life Sciences, Branford, CT, USA) and a 10-bp barcode sequence which is unique to each individual sample, and 806R attached to a Roche 454-B sequencing adapter. Polymerase Chain Reaction (PCR) amplification was performed in a volume of 50 μL reaction, each containing 5 μL of FastStart High Fidelity Reaction Buffer with 1.8 mM $MgCl_2$, 200μM dNTPs, 0.4 μM forward and reverse primers, 10-100 ng of DNA, and 2.5 U FastStart High Fidelity Enzyme Blend (Roche Diagnostics, Germany). The following thermal cycling program was used for PCR: 94 °C for 3 min for 1 cycle; 94 °C for 30 sec, 57 °C for 1 min, 72 °C for 2 min for 20 cycles; and a final extension at 72 °C for 2 min. PCR products were purified with the Qiagen PCR purification kit (Qiagen, Valencia, California, USA) and the Agencourt AMPure PCR purification system (Beckman Coulter, Danvers, Massachusetts, USA). The quality of each amplicon library was measured using the Agilent DNA 7500 kit with a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, USA). Equal molar of amplicons from each sample were pooled together. The pooled DNA was immobilized onto DNA

capture beads and amplified through emulsion PCR using the GS FLX titanium emPCR amplicon kit according to the manufacturer's protocols (454 Life Sciences, Branford, CT, USA). Sequencing of the PCR products was performed at the Center for Environmental Biotechnology at the University of Tennessee using a 454 Genome Sequencer FLX titanium platform (454 Life Sciences, Branford, CT, USA).

4.3.5. Bacteria isolation and identification by 16S rDNA sequencing

Bacteria strains were isolated and purified using R2A agar at 28°C. Bacterial 16S rRNA genes were amplified by bacterial universal primers 8F (5'-AGAGTTTGATCMTGGCTCAG-3') (Turner et al. 1999) and 907R (5'-CCGTCAATTCMTTTRAGTTT-3') (Lane 1991). Each PCR reaction mixture contained 0.4 μM of each primer, 200 μM dNTP, 2.5 U Ex Taq DNA polymerase, PCR buffer mix provided by the supplier of the Taq DNA polymerase (Takara, Madison, Wisconsin, USA), and 10 ng DNA template. PCR was performed with the following thermal cycling program: 94°C for 5 min, 30 cycles at 94°C for 1 min, 54°C for 1 min, and 72°C for 1 min, followed by a final extension at 72°C for 6 min. The amplified DNA products were purified using Qiagen PCR purification kit (Qiagen, Valencia, California, USA) and sequences using the reverse primer an ABI 3730xl DNA Analyzer (Applied Biosystems, Foster city, CA, USA).

4.3.6. Sequence analysis

The analysis of 16S rRNA gene sequences obtained from pyrosequencing were performed with MOTHUR v.1.23.1 (Schloss et al. 2009). Sequences were denoised, sorted and trimmed according to sample-specific barcodes. Dereplication, chimera check, low quality and short sequences (<100bp) removal were performed base on MOTHUR pipeline as described

previously (Huse et al. 2010). The non-redundant sequences were aligned against the SILVA database (http://www.arb-silva.de/). Operational taxonomic units (OTUs) were assigned with the average neighbor clustering algorithm and taxonomy was assigned using RDP naïve Bayesian rRNA classifier with a bootstrap cutoff of 80% (Wang et al. 2007). Hierarchical cluster analysis was based on family level bacterial community composition and performed using weighted Unifrac metrices. The neighbor joining phylogenetic trees for bacteria isolates were constructed by aligning with their closest relatives blast in NCBI database (http://www.ncbi.nlm.nih.gov/BLAST/) through ClustalX1.83 and clustering through Mega 5.0 (Tamura et al. 2011, Thompson et al. 1997).

Nucleotide sequence accession numbers

454 pyrosequencing raw data was submitted to NCBI GenBank Short Read Archive (SRA) with the accession number SRA051322.1.

4.4. Results and Discussion

4.4.1. Water quality changes through treatment, distribution and stagnation

The water quality changes during treatment, distribution and stagnation processes were shown in Table 4.1. The raw water used by WTPA had relatively higher turbidity, DOC, ions concentrations than the raw water used by WTPB. Consequently, the samples taken in the middle of treatment, as well as the tap water samples, received from WTPB had higher turbidity, DOC, ions concentrations. The pH values in WTPA and WTPB were controlled to 7.99 and 7.32, respectively, and hardly changed after distribution and stagnation. Coagulation and filtration

effectively removed the particles in river water and bring turbidity and DOC down to very low level in both water treatment plants. The turbidity was relative stable during the downstream treatment, distribution, and stagnation in both plants. The DOC concentrations also stayed in a low level. Sulfate and nitrate were far below the maximum contaminant levels (MCL) (250 mg/L and 10 mg/L, respectively) for drinking water regulated by the US Environmental Protection Agency (EPA). The free chlorine controlled by WTPA and WTPB were 2.24 mg/L and 2.73 mg/L, respectively. After distribution and stagnation, slightly decrease of free chlorine concentration were observed in both tap water and stagnant water. Heterotrophic plate count (HPC) was used to monitor the heterotrophic bacteria numbers living in the water after each treatment steps. HPC in the raw water used by both water treatment plants was about 10^6 CFU/100ml. After filtration average log removals for HPC were 1.4 and 5.5 in WTPA and WTPB, respectively. And then about 3.3 and 0.9 log removal were achieved after disinfection in WTPA and WTPB, respectively. Disinfection was the key step for bacteria removal in membrane filtration plant. In the conventional filtration plant a major decrease for HPC was observed after the filtration step. An increase of HPC was detected after distribution in the tap water supplied by WTPB. After stagnation obvious increase for HPC was also detected in the stagnant tap water supplied by both water treatment plants.

Table 4.1 Summary of water quality data during treatment, distribution and stagnation processes

Sample[a]	pH	Turbidity (NTU)	DOC (mg/L)	Free Cl_2 (mg/L)	Conductivity (uS/cm)	chloride (mg/L)	Sulfate (mg/L)	Nitrate (mg/L)	HPC (CFU/100ml)
AR	7.57	2.544 ± 0.049	1.01 ± 0.00	ND	198	2.28 ± 0.02	5.81 ± 0.01	1.07 ± 0.01	3500000 ± 707107
AM	7.76	0.014 ± 0.004	0.81 ± 0.05	ND	201	2.82 ± 0.01	5.67 ± 0.01	1.06 ± 0.01	147750 ± 7425
AF	7.99	0.015 ± 0.001	0.44 ± 0.01	2.24 ± 0.00	234	7.63 ± 0.13	5.20 ± 0.26	0.96 ± 0.00	80 ± 1
Afr	7.87	0.016 ± 0.001	0.54 ± 0.00	2.20 ± 0.02	210	7.78 ± 0.08	5.54 ± 0.01	1.10 ± 0.01	29 ± 0.00
Ast	7.85	0.018 ± 0.001	0.74 ± 0.00	1.93 ± 0.02	212	7.69 ± 0.00	6.05 ± 0.01	1.12 ± 0.00	110 ± 20
BR	7.89	9.379± 0.141	2.19 ± 0.01	ND	206	7.53 ± 0.05	11.51 ± 0.02	1.96 ± 0.00	6500000 ± 1343503
BS	7.20	0.034 ± 0.013	1.50 ± 0.11	ND	191	12.48 ± 0.09	22.54 ± 0.24	1.78 ± 0.07	19 ± 2
BF	7.32	0.037 ± 0.008	1.37 ± 0.02	2.73 ± 0.04	208	13.69 ± 0.10	22.52 ± 0.07	1.77 ± 0.01	1 ± 2
Bfr	7.21	0.038 ± 0.001	1.25 ± 0.02	2.14 ± 0.02	263	17.91 ± 0.07	29.22 ± 0.13	1.00 ± 0.00	138 ± 33
Bst	7.25	0.039 ± 0.001	1.29 ± 0.06	1.93 ± 0.04	256	17.15 ± 0.01	25.91 ± 0.15	0.84 ± 0.02	1598 ± 32

[a]AR, AM, AF, Afr and Ast represented water samples of water treatment plant A taken from raw water, after membrane filtration, finished water, after distribution and after stagnation; BR, BS, BF, Bfr and Bst represented water samples of water treatment plant B taken from raw water, after sand filtration, finished water, after distribution and after stagnation.

4.4.2. Overview of bacterial community composition in drinking water treatment and supply systems

One of the main objectives of drinking water treatment is to remove microorganisms in the raw water. In order to monitor the bacterial community changes after each treatment step, 16S rRNA gene amplicon libraries were constructed from raw water, filtrated water, chlorinated water, distributed tap water and stagnant tap water, respectively. Bacterial communities were analyzed through pyrosequencing of the 16S rRNA gene amplicons. Total of 3051 and 4343 sequences, with an average length of 250bp, were obtained from membrane filtration plant (WTPA) and conventional filtration plant (WTPB), respectively. With a cut off value of 80%, these sequences were assigned to total of 8 and 14 bacterial phyla for samples from WTPA and WTPB, respectively. Bacterial members in *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, and *Firmicutes* were found dominated in both water treatment plants (Figure 4.1). *Proteobacteria*

was the most dominant phylum in both water treatment plants, with the relative abundance ranged from 52.3% to 99.4% in water samples from WTPA, and 49.0% to 99.0% in WTPB water samples, respectively.

Obvious bacterial community changes were observed at phylum level, after each treatment step. The two source water had similar bacterial communities although they were from different rivers. As shown in Figure 4.1 the dominant bacteria were distributed in *Alphaproteobacteria* (6.9% and 13.1%), *Betaproteobacteria* (44.4% and 37.5%), *Gammaproteobacteria* (0.6% and 9.3%), *Bacteroidetes* (31.1% and 16.8%), and *Actinobacteria* (15.4% and 12.1%) for WTPA and WTPB source water. After membrane filtration only two major bacterial populations *Betaproteobacteria* (48.6%) and *Gammaproteobacteria* (48.7%) were found in the permeate water of WTPA (Figure 4.1A). After disinfection treatment, the relative abundance of G*ammaproteobacteria* decreased to 3.9%. *Betaproteobacteria* (66.6%), *Alphaproteobacteria* (17.76%) and *Firmicutes* (8.6%) became the dominant bacterial populations present in the finished water of WTPA. After distributed to customer's tap, the relative abundance of *Alphaproteobacteria* and *Firmicutes* in the tap water was decreased to 1.4% and 0.2%, respectively. The *Betaproteobacteria* (88.4%) and *Gammaproteobacteria* (9.4%) were found dominated in the distributed tap water. However bacterial community showed different pattern after stagnation. The relative abundance of *Alphaproteobacteria* and *Firmicutes* increased to 68.4% and 11.4% present as predominant populations. *Betaproteobacteria* (10.4%) and *Gammaproteobacteria* (5.2%) were also persistent in the stagnant water with relatively high abundance. During the treatment, distribution, and stagnation processes in WTPA, *Betaproteobacteria* was the dominant populations persistent after each step from the raw river water to the tap water at the costumer's end. Substantial changes in bacterial community

composition were also observed in the conventional filtration plant along the treatment and distribution pipeline (Figure 4. 1B). After processed through coagulation, flocculation and sand filtration, the relative abundance of *Bacteroidetes* and *Actinobacteria* decreased below 0.4%. *Alphaproteobacteria* (34.3%), *Betaproteobacteria* (37.8%) as well as *Gammaproteobacteria* (6.8%) and *Firmicutes* (15.8%) became the dominant bacterial populations in the filtrated water. After chlorination, the relative abundance of *Alphaproteobacteria, Gammaproteobacteria* and *Firmicutes* decreased below 0.4%. *Betaproteobacteria* (95.8%) was the only dominant bacterial population observed in the finished drinking water of WTPB. Passed through the distribution system, high abundance of *Bacteroidetes* (43.5%) and *Betaproteobacteria* (37.5%), as well as *Alphaproteobacteria* (7.5%), *Gammaproteobacteria* (2.8%) and *Actinobacteria* (7.5%), were detected in the distributed tap water produced by WTPB. After stagnation, *Alphaproteobacteria, Actinobacteria,* and *Firmicutes* increased to 77.9%, 9.8%, and 8.6% respectively. *Betaproteobacteria* (1.7%) and *Bacteroidetes* (0.5%) were present with relatively low abundance in the stagnant water. Similarly as what we observed in the membrane filtration plant, *Betaproteobacteria* was the major bacterial populations persistent during the whole treatment and distribution process for drinking water supply.

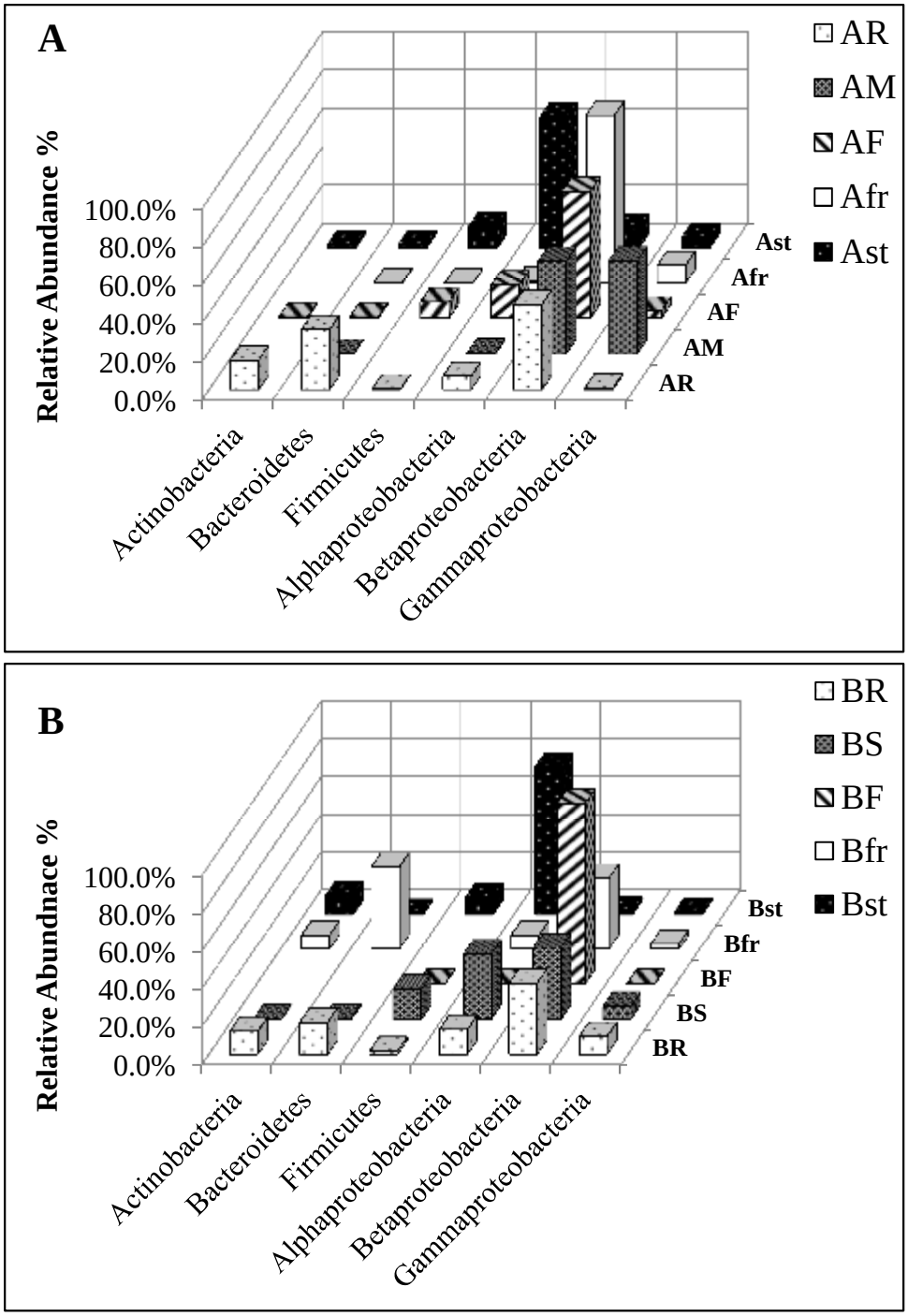

Figure 4.1 Bacterial community composition identified by 16S rRNA gene pyrosequencing from water treatment plant A (A) and water treatment plant B (B). Bacteria phyla and *Proteobacteria* classes with relative abundance > 1% are shown.

In summary, varieties of bacteria including *Alphaproteobacteria*, *Betaproteobacteria*, *Gammaproteobacteria,* and *Firmicutes* cannot be eliminated through conventional coagulation and sand filtration processes. Membrane filtration was able to remove a large variety of bacteria populations; however, was less effective in removing bacteria belonged to the class of *Betaproteobacteria* and *Gammaproteobacteria*. In both water treatment plants *Betaproteobacteria* was the dominant populations survived after each treatment step, including filtration and disinfection, and persistent in the distribution systems. After stagnation the increase of *Alphaproteobacteria* and *Firmicutes* were detected in the stagnant water supplied by both water treatment plants.

4.4.2.1. Bacterial population dynamics during drinking water treatment, distribution, and stagnation processes.

To track the bacterial community fingerprints changes during drinking water treatment, distribution and stagnation processes, phylogenetic analyses were performed down to the family level and genus level when the taxonomic assignment allows. Around 95.0% and 79.5% sequences from water samples of WTPA and WTPB were classifiable at family level. Total of 47 and 63 bacterial families were identified in water samples from WTPA and WTPB respectively. The bacterial community composition with abundant bacterial families (>2%) were shown in Figure 4.2 and 3.

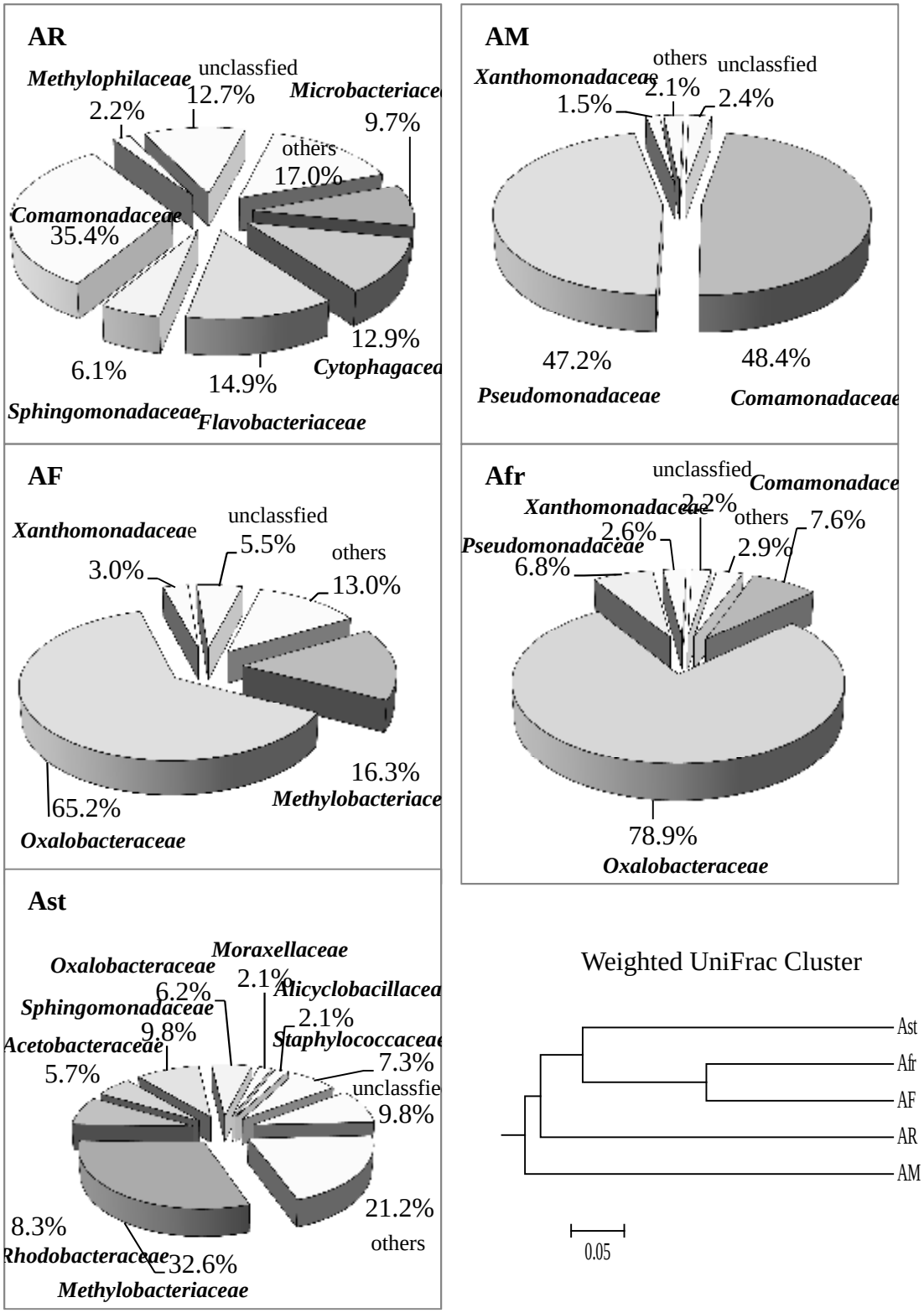

Figure 4.2 Bacterial community composition identified by 16S rRNA gene pyrosequencing from water treatment plant A. Abundant families with relative abundance > 2% were shown.

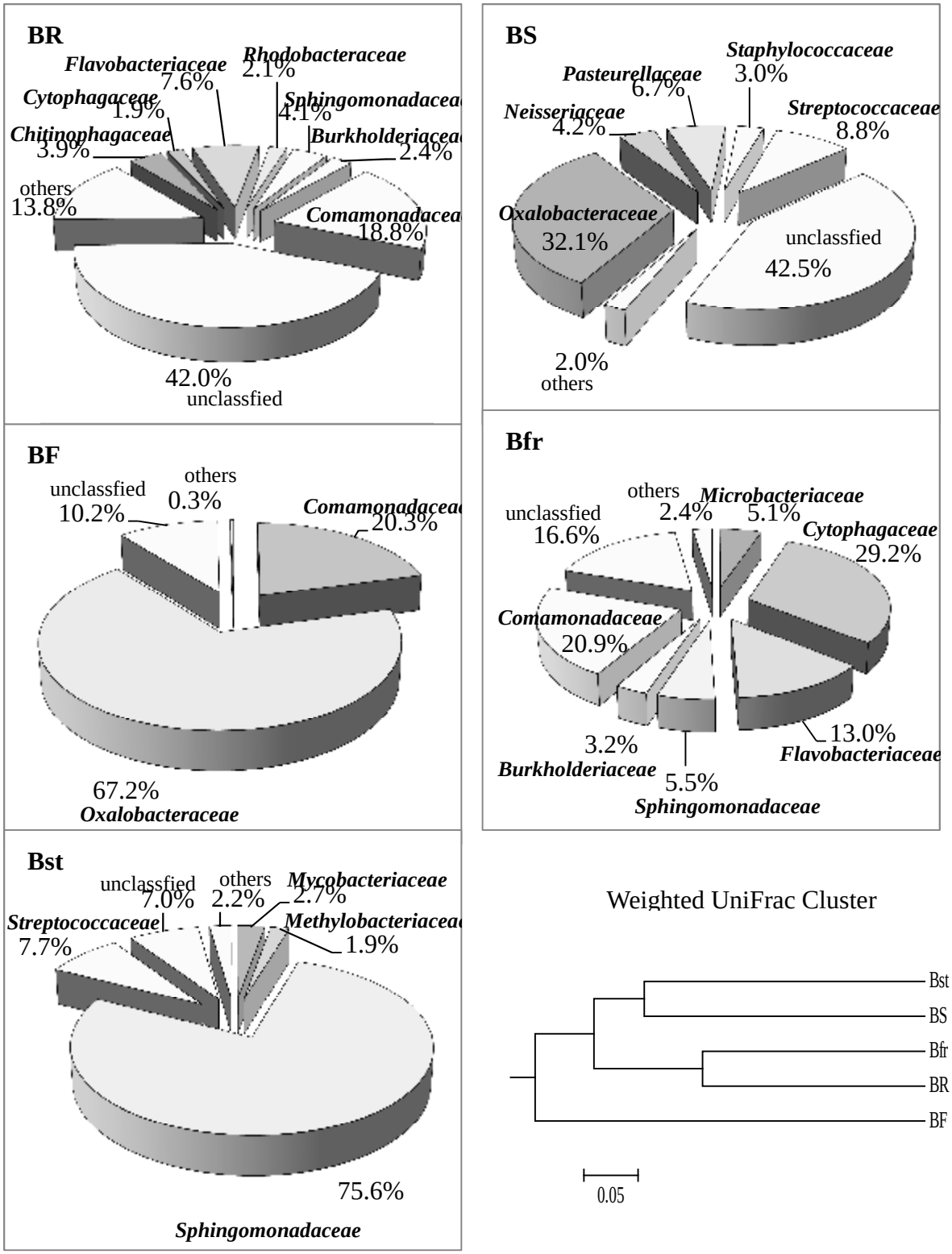

Figure 4.3 Bacterial community composition identified by 16S rRNA gene pyrosequencing from water treatment plant B. Abundant families with relative abundance > 2% were shown.

4.4.2.2. Bacterial population change after filtration

Bacterial community composition was greatly changed after the coagulation, flocculation treatment, and membrane sieving (Figure 4.2). Most bacterial families in the raw water could not be detected in the membrane filtered water, except for *Comamonadaceae* which became the most dominant bacteria family in the permeated water. Another bacteria family *Pseudomonadaceae* (47.2%) survived the membrane sieving and existed in the permeated water with similar relative abundance as *Comamonadaceae*. *Comamonadaceae* and *Pseudomonadaceae* were affiliated to the class of *Betaproteobacteria* and *Gammaproteobacteria*, respectively. Bacterial sequences from these two families in the permeated water were further classified down to the genus of *Delftia* (45.1%) and *Pseudomonas* (40.5%), respectively, suggesting that ultrafiltration membranes may have less efficiency for the removal of bacteria from these to genera. However, we still cannot concluded that bacteria from these two genera were able to pass through membrane sieves since the analysis was based DNA analysis both living cell and cell debris could be detected. Some species of *Delftia* were potential opportunistic pathogens and reported in contaminated tap water (Jurado et al. 2002). *Pseudomonas* was observed from both membrane tank and membrane biofilm in a pilot membrane filtration plant (Kwon et al. 2011).

Conventional filtration process also greatly changed the bacterial community composition, as seen by the low similarity between the surface water and the water after filtration (Figure 4.3). However, unlike the membrane permeated water, varieties of bacteria affiliated to *Alphaproteobacteria Betaproteobacteria* Gamma*proteobacteria* and *Firmicutes* were detected in the sand filtration water. Most *Alphaproteobacteria* sequences in the filtrated water couldn’t be classified at deeper taxonomic level. A small amount of *Alphaproteobacteria* sequences were recognized as *Sphingomonadaceae* (0.48%). The sequences affiliated to *Betaproteobacteria* were

assigned to the family of *Oxalobacteraceae* (32.1%) and *Neisseriaceae* (4.2%). *Oxalobacteraceae* was assigned to the genus of *Undibacterium* (26.7%) and *Massilia* (3.9%). *Neisseriaceae* (4.2%) was assigned to the genus of *Neisseria* (4.0%). Most *Gammaproteobacteria* sequences were assigned to the family of *Pasteurellaceae* (6.7%) and couldn't be classified at genus level. *Firmicutes* was mainly composed of *Streptococcaceae* (8.8%) and *Staphylococcaceae* (3.0%) which were assigned to the genus of *Streptococcus* (8.8%) and *Staphylococcus* (2.9%), respectively.

Different from our results Eichler found that the influence of unspecific coagulation, flocculation and sand filtration treatment on bacteria community structure was very small (Eichler et al. 2006). Similar to our results, Pinto et al found that filtration step play an important role in shaping the bacterial community in drinking water distribution system (Pinto et al. 2012). The phylotypes detected by Eichler from the raw and the sand filtered sample were also very different from the bacteria genera we observed. The great variances of bacteria communities may lead to the difference in treatment efficiency.

4.4.2.3. Bacterial population change after disinfection

Chlorination was a main step to deactivate microorganisms and the last barrier to these microorganisms. Consequently, substantial bacterial population shift prior to and after chlorination were observed in both water treatment plants. *Comamonadaceae*, the most abundant population in the membrane permeated water was not detectable after chlorination. *Pseudomonadaceae* another abundant family was detected with low abundance (0.8%) in the finished water. *Xanthomonadaceae* was observed in water before and after chlorination with low relative abundance (1.5% and 3.0%, respectively). Consistent with permeate water, bacteria sequences affiliated to the family of *Pseudomonadaceae* and *Xanthomonadaceae* in the finished

water were also classified into the genus of *Pseudomonas* and *Stenotrophomonas*, respectively. Different from the permeate water, the dominate bacteria in the finished water belonged to the family of *Oxalobacteraceae* (65.2%) and *Methylobacteriaceae* (16.3%), which affiliated to *Alphaproteobacteria* and *Betaproteobacteria*, respectively. A small amount of *Oxalobacteraceae* was assigned to the genus of *Massilia* (9.7%) and *Naxibacter* (5.8%) and a large amount of *Oxalobacteraceae* was not classified to a deeper level. Most *Methylobacteriaceae* sequences found in the finished water were further assigned to the genus of *Methylobacterium* (16.0%). Therefore bacteria members from the *Delftia* genus may be sensitive to chlorine disinfection. *Pseudomonas*, *Stenotrophomonas*, *Methylobacterium*, *Massilia*, *Naxibacter,* and the unclassified *Oxalobacteraceae* might be potential chlorine resistant population that can survive disinfection step.

In WTPB the disinfection process also eliminated most bacterial families from the sand filtrated water, except for *Oxalobacteraceae* which became the most abundant population (67.2%) in the finished water. Another abundant population persistent in the finished water belonged to the family of *Comamonadaceae* (20.3%). Most bacterial sequences affiliated to the two dominant families were assigned to the genus of *Undibacterium* (67.1%) and *Acidovorax* (13.9%), respectively. The origin of *Acidovorax* can be traced back to the raw river water. The *Undibacterium* was also detected in the sand filtered water but with lower relative abundance. The increased relative abundance of *Undibacterium* and *Acidovorax* after chlorination treatment indicated their chlorine tolerant ability. *Undibacterium* is newly observed genus added to the family *Oxalobacteraceae* in recent years. It has previously been isolated from drinking water and bottle water with resistant features to some antibiotics (Falcone-Dias et al. 2012, Kämpfer et al. 2007). *Acidovorax* was also recently reported to be persistent in different treatment stages during

conventional filtration treatment processes (Pinto et al. 2012).

Oxalobacteraceae was found dominated in the finished water of both WTPA and WTPB. Similar with our results *Oxalobacteraceae* as well as *Comamonadaceae* were also found predominant in the finished water of three conventional filtration water treatment plants using different source water (Poitelon et al. 2009). However, in Poitelon's results *Oxalobacteraceae* and *Comamonadaceae* were mostly classified into the genus of *Massilia* and *Polaromonas,* respectively. *Stenotrophomonas* was also detected by Poitelon in their finished water using conventional treatment process but with low abundance level. The consistency in family level but diversity in genus level suggesting that chlorine resistant bacteria population may distribute in variety of bacterial genera but belonged to several consistent core bacterial families. Unlike our results that *Betaproteobacteria* showed chlorine resistance potential in than other bacteria phyla, some study conclude that *Betaproteobacteria* was more sensitive to chlorine than *Alphaproteobacteria*, others found that *Gammaproteobacteria* and *Detaproteobacteria* were tolerant to chlorine (Kormas et al. 2010, Poitelon et al. 2010, Williams et al. 2004). The differences of bacteria community in their source river may be another reasonable explanation for the discrepancy of bacteria communities in finished water reported in different studies. In addition, the regulated chlorine concentration used by different water treatment plants was significantly different from each other, varied from below 1 mg/L to about 3 mg/L. Different bacteria may have different tolerant level to chlorine. This may also lead to the inconsistent results for chlorine resistance bacteria populations. Similar to our results Poitelon also concluded that chlorination played an important role in shaping bacterial community released to drinking water distribution systems (Poitelon et al. 2010).

4.4.2.4. Bacterial population change after distribution

After disinfection the drinking water is sent to the customer's house though water distribution pipeline. Most detected bacteria families and genera in the finished water were mainly originated from finished water and membrane permeated water (Figure 4.2 AM, AF and Afr). Similar to the finished water, *Oxalobacteraceae* (78.9%) was still predominate population in the fresh distributed tap water from WTPA. Genus level bacterial composition is also similar for the two water samples, as a small amount of *Oxalobacteraceae* was identified as *Massilia* (12.8%) and *Naxibacter* (7.1%) and the rest was unclassified. *Xanthomonadaceae* (2.6%) in the fresh distributed tap water also classified into the genus of *Stenotrophomonas* (2.6%) with similar relative abundance as what we observed in the finished water. However the second dominated population *Methylobacteriaceae*, which identified as *Methylobacterium*, appeared as a minor population (0.2%) after distributed through the water distribution pipelines. *Comamonadaceae* (7.6%) and *Pseudomonadaceae* (6.8%) in the fresh tap water were identified as *Delftia* (6.2%) and *Pseudomonas* (6.0%) which existed in the membrane permeate water as the two dominate populations.

However substantial bacterial population changes were observed between the finished water and the water after distributed through WTPB's pipelines. The largest population in the finished water, *Oxalobacteraceae*, appeared as a minor population (0.4%) in distributed tap water. The *Oxalobacteraceae* sequences in the distributed water were classified in to the genus of *Massilia* (0.4%) instead of *Undibacterium* which dominated in both sand filtrated water and finished water, but not detected in the distributed water. *Comamonadaceae* in both finished water and distributed water had similar relative abundance (20.3% and 20.9%). However the most *Comamonadaceae* sequences in finished water were assigned to the genus of *Acidovorax*. In

distributed water only a small amount of *Acidovorax* (0.4%) were observed, while most *Comamonadaceae* sequences were unclassified. *Bacteroidetes* from the family of *Cytophagaceae* (29.2%) and *Flavobacteriadeae* (13.0%) became dominant population in the water after distribution. These two bacteria families were not detected after sand filtration and disinfection but detected in the raw river water with the relative abundance of 1.9% and 7.6%, respectively. *Flavobacteriadeae* associated sequences appeared in both river water and distributed tap water were assigned to the genus of *Flavobacterium*. There were some other families also appeared with relative high abundance after distribution process such as *Microbacteriaceae* affiliated to the phylum of *Actinobacteria*, which was not observed either after sand filtration or disinfection process but detected in the raw water with low relative abundance (0.6%). *Sphingomonadaceae* (5.5%) and *Burkholderiaceae* (3.2%), affiliated to the class of *Alphaproteobacteria* and *Betaproteobacteria*, in the distributed tap water were also detected during the treatment process and raw water. Small amount of *Mycobacterium* (0.7% in BR and 1.2% in Bfr) which affiliated to the family of *Mycobacteriaceae* was observed on both river water and tap water samples.

For WTPA the origin of bacteria in tap water can be traced back to the water treatment plant. However, bacteria tap water from WTPB was more diverse than the finished water. The newly appeared bacteria populations can be detected in the source river but not in filtered and chlorinated finished water. Biofilm inside water distribution networks could be another source for tap water bacteria. Since the only origin of biofilm bacteria should be the river water, the aged biofilm fall off from water distribution pipelines could be one reasonable origin for the newly appeared populations in the tap water.

Massilia is widely distributed in natural environment as soil and air, has been isolated

recently from the drinking water showing the tendency to form pellicles in liquid medium (Gallego et al. 2006b). *Massilia* was also found to be the most predominant population in the biofilm formed in drinking water distribution system (Liu et al. 2012, White et al. 2011). *Naxibacter* is new member added in to the family of O*xalobacteraceae in recent years. Naxibacter* isolates were reported distributed in soil and but haven't be observed in water sample (Weon et al. 2010, Xu et al. 2005). *Methylobacterium* is widely distributed in aquatic environments and have been isolated from bulk water and biofilm in water distribution system (Hiraishi et al. 1995). Recent studies on the physiology of *Methylobacterium* isolates from drinking water showed that these bacteria are capable of forming biofilm and using a diverse group of carbon substrates including C1 compounds (Gallego et al. 2005, 2006a, Simoes et al. 2010). *Flavobacterium* is frequently observed in aquatic environments (Zwart et al. 2002). It was also detected in finished water through 16S rDNA based analysis and isolated from disinfected drinking water with a high degree of resistance to chlorine (Poitelon et al. 2010, Stewart et al. 1990). *Mycobacterium* has been found in finished water, tap water and biofilm attached inside the water distribution pipes (Liu et al. 2012, Poitelon et al. 2010). Besides, some *Mycobacterium* were found to be resistant to chlorine treatment, some species are well known human pathogens (Le Dantec et al. 2002a, b). The bacteria species exist in tap water either have chlorine resistance or have the ability to form biofilm and capable of using low concentrate organics matters in drinking water as food source.

4.4.2.5. Bacterial population change after stagnation

The family level fingerprints also showed substantial differences between the bacterial community of fresh distributed tap water and stagnate tap water. After stagnation the most dominant population in the fresh distributed water of WTPA, *Oxalobacteraceae* decreased to

6.2%. The bacteria genera associated to *Oxalobacteraceae* in the stagnant water were *Massilia* (1.6%), *Janthinobacterium* (1.0%), and *Naxibacter* (0.5%). The second abundant population in the fresh distributed water, *Comamonadaceae* was decreased below 1.6% after stagnation and assigned to the genus of *Acidovorax* instead of *Delftia*. The other two abundant populations in the fresh water, *Pseudomonadaceae* and *Xanthomonadaceae* decreased below 1.0%. The *Betaproteobacteria* which dominated in the fresh tap water greatly decreased. Instead, family members affiliated to *Alphaproteobacteria* and *Firmicutes* became the dominant population after stagnation. The *Alphaproteobacteria* existed in the stagnant tap water of WTPA was mainly associated with the family of *Methylobacteriaceae* (32.6%), *Sphingomonadaceae* (9.8%), *Rhodobacteraceae* (8.3%) and *Acetobacteraceae* (5.7%). The *Firmicutes* affiliated sequences in the stagnant tap water were assigned to the family of *Staphylococcaceae* (7.3%) and *Alicyclobacillaceae* (2.1%). Sequences affiliate to the family of *Methylobacteriaceae*, *Rhodobacteraceae*, *Sphingomonadaceae*, *Staphylococcaceae* and *Alicyclobacillaceae* were classified to the genus of *Methylobacterium* (32.6%), *Paracoccus* (1.6%), *Sphingomonas* (8.3%), *Staphylococcus* (7.3%), and *Tumebacillus* (2.1%), respectively.

Similar to what we observed in WTPA, in WTPB after stagnation bacteria families in *Betaproteobacteria* greatly decreased, instead bacteria families in *Alphaproteobacteria* and *Firmicutes* greatly increased and became the dominant population. The *Alphaproteobacteria* sequences increased in the stagnant tap water of WTPB was assigned to the family of *Sphingomonadaceae* (75.6%) and *Methylobacteriaceae* (1.9%). Sequences affiliated with these bacterial families were further classified into the genus of *Sphingomonas* (72.4%) and *Methylobacterium* (1.9%). The *Firmicutes* sequences were assigned to the family of *Streptococcaceae* (7.7%) and genus of *Streptococcus* (7.7%). bacteria families affiliated to the

phylum of *Actinobacteria* also changed after stagnation. After stagnation, *Microbacteriaceae* was not detected while *Mycobacteraceae* increased from 1.2% in the fresh distributed water to 2.7% in the stagnated water. *Mycobacteraceae* was further classified as *Mycobacterium*, which was also detected in the raw water with relative low abundance (0.7%). However, the most dominated population appeared in fresh tap water, *Bacteroidetes* decreased below 0.5% after stagnation.

The increase of bacteria number and shift of bacteria population were also observed by other studies (Lautenschlager et al. 2010, Pepper et al. 2004). Pepper reported a shift from Gram-negative population to a higher percentage of Gram-positive populations using culture based method. Variety of Gram-positive bacteria was also observed in our study such as *Staphylococcus*, *Streptococcus*, and *Tumebacillus*. *Tumebacillus* as a member from the *Firmicutes* phylum has been observed in surface water (Liu et al. 2011). Some pathogens from the genus of *Staphylococcus*, *Streptococcus*, and *Mycobacterium* were also observed after stagnation. However their biological activity requires further study. We also found the increase of *Sphingomonas* and *Methylobacterium* in taps from both water treatment plants after stagnation. Both of them are Gram negative bacteria. Bacterial members from the genus *Sphingomonas* and *Methylobacterium* were frequently isolated from drinking water and biofilm in drinking water distribution system (Koskinen et al. 2000, Srinivasan et al. 2008). Some members from the genus of *Paracoccus* were famous denitrifies. There are also aerobic *Paracoccus* distributed in natural environment such as soil and ocean as well as activate sludge (Dastager et al. 2011, Sheu et al. 2011, Sun et al. 2012). The changes in bacterial community may be caused by either the regrowth or biofilm detachment. Therefore both the differences in the source water communities and study method may cause the different observations in bacteria population change.

4.4.3. Core population change during drinking water treatment and distribution

Taxonomic analysis revealed a broad range of bacteria widely distributed in drinking water samples. Some resilient populations were observed with high frequency during water treatment processes. There were 8 bacteria families found persistent in at least 3 water samples in WTPA and 6 bacteria families detected in more than 3 samples in WTPB. The shared families and their dynamic changes during treatment process were shown in Figure 4.4. Five of those families were found in both water treatment plants. Bacteria members from the family of *Sphingomonadaceae*, *Burkholderiaceae*, *Comamonadaceae*, *Oxalobacteraceae,* and *Xanthomonadaceae* were the core resilient populations have high chance to survive treatment and distribution process. *Comamonadaceae* and *Oxalobacteraceae* were persistent with high relative abundance during chlorination and distribution. *Xanthomonadaceae* always appeared with low abundance. The relative abundance of *Sphingomonadaceae* greatly increased in tap water from both water treatment plants after stagnation. There were four potential pathogens detected in chlorinated water with highest relative abundance usually appeared after stagnation: *Mycobacterium* (1.2% in Bfr and 2.7% in Bst), *Stenotrophomonas* (3.0% in AF and 2.6% in Afr), *Staphylococcus* (0.8% in AF, 0.1% in BF, 0.2% in Afr, and 7.3% in Ast), and *Streptococcus* (0.1% in BF and 7.7% in Bst).

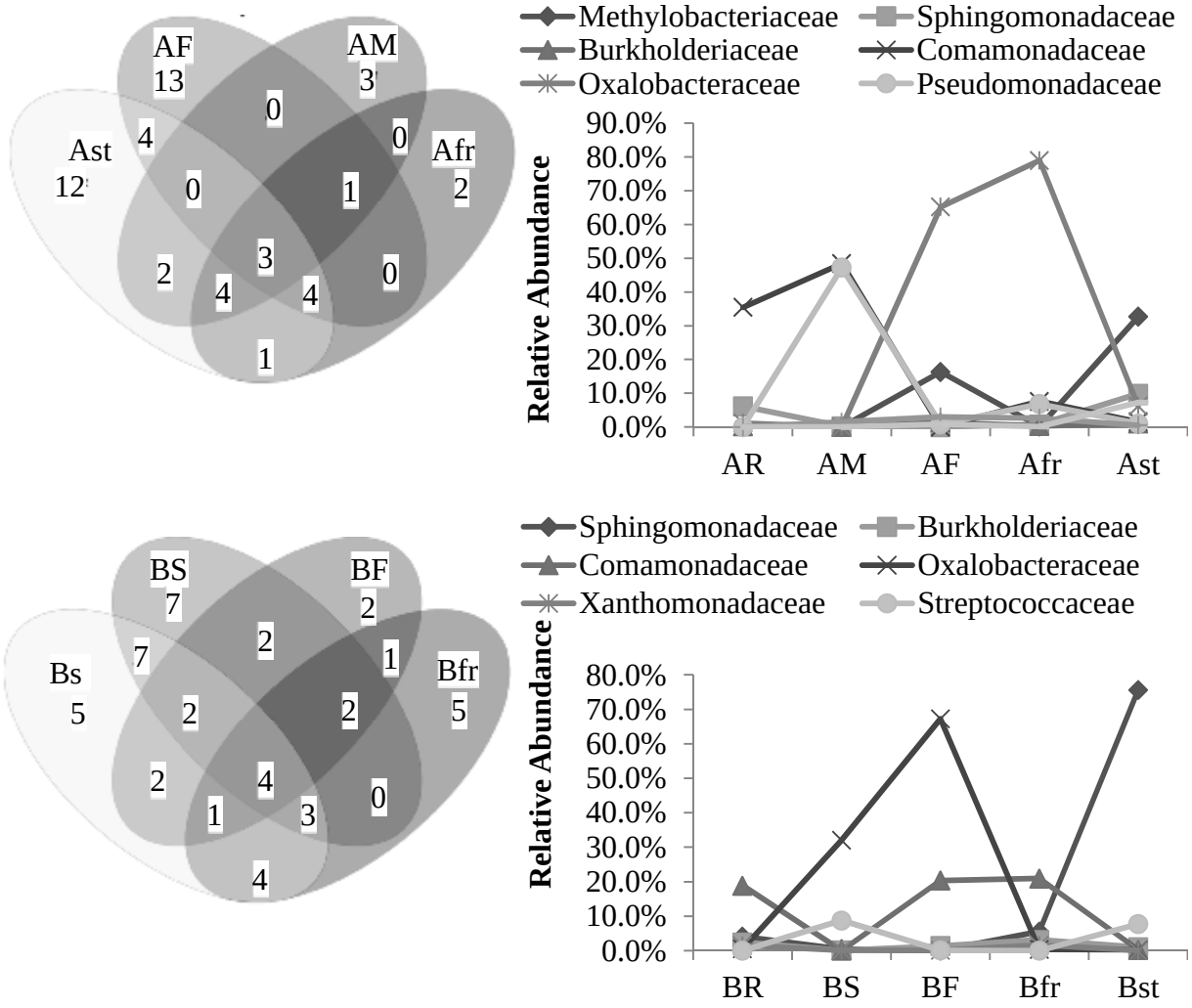

Figure 4.4 Venn diagram of shared bacterial families and the dynamic changes of some core bacterial families.

4.4.4. Bacteria isolated after each treatment step

Bacteria detected using 16 S rDNA analyses may include both live and dead cells, which may subsequently bias the community structure. However bacteria isolates can provide solid evidence for the presence of living cells in drinking water. To further confirm the active bacteria may exist during water treatment processes bacteria were isolated using R2A agar from water sample taken after each treatment step. The phylogenetic tree of bacteria isolates were shown in Figure 4.5. Bacteria strains from ***Pseudomonas*** were isolated from the membrane permeated water sample suggesting the ineffectiveness of membrane ultrafiltration in ***Pseudomonas*** removal. However no ***Delftia*** strains were isolated after membrane filtration. One possible reason is the DNA level detection also included cell debris with DNA remaining inside as well as active bacteria cells. Another reason for the failure of isolation may be cause by the limited culturing ingredients in the media. Bacteria strains belong to ***Pseudomonas*** and ***Stenotrophomonas*** were also isolated from corresponding finished water samples. The presence of ***Pseudomonas***, ***Stenotrophomonas,*** and ***Naxibacter*** were also verified by the tap water isolation. The occurrence of ***Sphingomonas*** and ***Paracoccus*** during stagnation in drinking water was confirmed by the corresponding isolates. The appearance of ***Mycobacterium***, ***Methylobacterium,*** and ***Sphingomonas*** in tap water and stagnated tap water received from WTPB was verified by bacteria isolates.

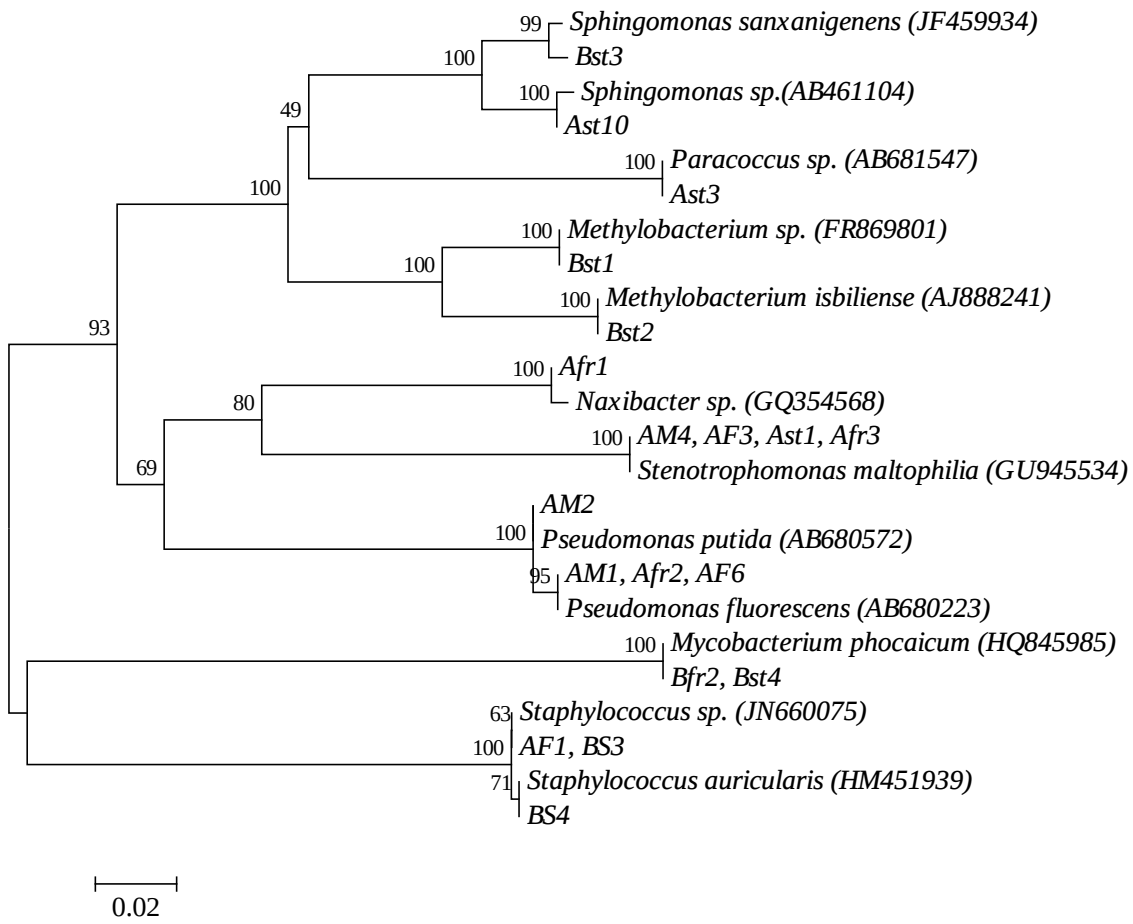

Figure 4.5 Neighbor-joining phylogenetic tree of 16S rDNA sequences from representative isolates in two water treatment plants. The numbers at the nodes indicate the percentages of occurrence in 1000 bootstrapped.

4.5. Conclusions

Substantial differences were observed after each treatment and distribution steps for both treatment plants. *Betaproteobacteria* survived each treatment steps and *Alphaproteobacteria* and *Firmicutes* increased after stagnation. Membrane filtration removed a large variety of bacteria populations; however, was less effective in removing bacteria from the genus of *Delftia* and *Pseudomonas*. For plant A chlorine disinfection was the key step for bacteria removal, subsequently playing an important role in shaping bacteria community structure in tap water. For plant B distribution system greatly affect the tap water bacterial community structure. After stagnation *Methylobacterium, Sphingomonas* increased in tap water from both water treatment plants. Core bacterial family detected with high frequency in multiple treatment steps in both water treatment plants: *Sphingomonadaceae*, *Burkholderiaceae*, *Comamonadaceae*, *Oxalobacteraceae,* and *Xanthomonadaceae.*

4.6. References

APHA (2005) Standard methods for the examination of water and wastewater., American Public Health Association, Washington, D.C.

Craun, G.F., Brunkard, J.M., Yoder, J.S., Roberts, V.A., Carpenter, J., Wade, T., Calderon, R.L., Roberts, J.M., Beach, M.J. and Roy, S.L. (2010) Causes of Outbreaks Associated with Drinking Water in the United States from 1971 to 2006. Clinical Microbiology Reviews 23(3), 507-528.

Dastager, S., Deepa, C., Li, W.-J., Tang, S.-K. and Pandey, A. (2011) Paracoccus niistensis sp. nov., isolated from forest soil, India. Antonie van Leeuwenhoek 99(3), 501-506.

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.

Falcone-Dias, M.F., Vaz-Moreira, I. and Manaia, C.M. (2012) Bottled mineral water as a potential source of antibiotic resistant bacteria. Water Research 46(11), 3612-3622.

Gallego, V., García, M.T. and Ventosa, A. (2005) *Methylobacterium hispanicum* sp. nov. and *Methylobacterium aquaticum* sp. nov., isolated from drinking water. International Journal of Systematic and Evolutionary Microbiology 55(1), 281-287.

Gallego, V., García, M.T. and Ventosa, A. (2006a) *Methylobacterium adhaesivum* sp. nov., a methylotrophic bacterium isolated from drinking water. International Journal of Systematic and Evolutionary Microbiology 56(2), 339-342.

Gallego, V., Sanchez-Porro, C., Garcia, M.T. and Ventosa, A. (2006b) Massilia aurea sp nov., isolated from drinking water. International Journal of Systematic and Evolutionary Microbiology 56, 2449-2453.

Henne, K., Kahlisch, L., Brettar, I. and Hofle, 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.

Hill, V.R., Kahler, A.M., Jothikumar, N., Johnson, T.B., Hahn, D. and Cromeans, T.L. (2007) Multistate evaluation of an ultrafiltration-based procedure for simultaneous recovery of enteric microbes in 100-liter tap water samples (vol 73, pg 4218, 2007). Applied and Environmental Microbiology 73(19), 6327-6327.

Hiraishi, A., Furuhata, K., Matsumoto, A., Koike, K.A., Fukuyama, M. and Tabuchi, K. (1995) Phenotypic and genetic diversity of chlorine-resistant *Methylobacterium* strains isolated from various environments. Applied and Environmental Microbiology 61(6), 2099-2107.

Ho, L., Braun, K., Fabris, R., Hoefel, D., Morran, J., Monis, P. and Drikas, M. (2012) Comparison of drinking water treatment process streams for optimal bacteriological water quality. Water Research 46(12), 3934-3942.

Huse, S.M., Welch, D.M., Morrison, H.G. and Sogin, M.L. (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environmental Microbiology 12(7), 1889-1898.

Jurado, V., Ortiz-Martinez, A., Gonzalez-delValle, M., Hermosin, B. and Saiz-Jimenez, C. (2002) Holy water fonts are reservoirs of pathogenic bacteria. Environmental Microbiology 4(10), 617-620.

Kämpfer, P., Rosselló-Mora, R., Hermansson, M., Persson, F., Huber, B., Falsen, E. and Busse, H.-J. (2007) Undibacterium pigrum gen. nov., sp. nov., isolated from drinking water. International Journal of Systematic and Evolutionary Microbiology 57(7), 1510-1515.

Kormas, K.A., Neofitou, C., Pachiadaki, M. and Koufostathi, E. (2010) Changes of the bacterial assemblages throughout an urban drinking water distribution system. Environmental Monitoring and Assessment 165(1-4), 27-38.

Koskinen, R., Ali-Vehmas, T., Kämpfer, P., Laurikkala, M., Tsitko, I., Kostyal, E., Atroshi, F. and Salkinoja-Salonen, M. (2000) Characterization of Sphingomonas isolates from Finnish and Swedish drinking water distribution systems. Journal of Applied Microbiology 89(4), 687-696.

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.

Lane, D.J. (1991) Nucleic acid techniques in bacterial systematics. Stackebrandt, E. and Goodfellow, M. (eds), pp. 115-175 John Wiley & Sons, New York, N.Y.

Lautenschlager, K., Boon, N., Wang, Y.Y., Egli, T. and Hammes, F. (2010) Overnight stagnation of drinking water in household taps induces microbial growth and changes in community composition. Water Research 44(17), 4868-4877.

Le Dantec, C., Duguet, J.-P., Montiel, A., Dumoutier, N., Dubrou, S. and Vincent, V. (2002a) Chlorine Disinfection of Atypical Mycobacteria Isolated from a Water Distribution System. Applied and Environmental Microbiology 68(3), 1025-1032.

Le Dantec, C., Duguet, J.-P., Montiel, A., Dumoutier, N., Dubrou, S. and Vincent, V. (2002b) Occurrence of Mycobacteria in Water Treatment Lines and in Water Distribution Systems. Applied and Environmental Microbiology 68(11), 5318-5325.

Li, D., Li, Z., Yu, J., Cao, N., Liu, R. and Yang, M. (2010) Characterization of bacterial community structure in a drinking water distribution system during an occurrence of red water. Appl. Environ. Microbiol. 76(21), 7171-7180.

Liu, R., Yu, Z., Zhang, H., Yang, M., Shi, B. and Liu, X. (2012) Diversity of bacteria and mycobacteria in biofilms of two urban drinking water distribution systems. Canadian Journal of Microbiology 58(3), 261-270.

Liu, Y., Yao, T., Jiao, N., Tian, L., Hu, A., Yu, W. and Li, S. (2011) Microbial diversity in the snow, a moraine lake and a stream in Himalayan glacier. Extremophiles 15(3), 411-421.

Mull, B. and Hill, V.R. (2009) Recovery and Detection of Escherichia coli O157:H7 in Surface Water, Using Ultrafiltration and Real-Time PCR. Applied and Environmental Microbiology 75(11), 3593-3597.

Nikkari, S., Lopez, F.A., Lepp, P.W., Cieslak, P.R., Ladd-Wilson, S., Passaro, D., Danila, R. and Relman, D.A. (2002) Broad-range bacterial detection and the analysis of unexplained death and critical illness. Emerging Infectious Diseases 8(2), 188-194.

Pepper, I.L., Rusin, P., Quintanar, D.R., Haney, C., Josephson, K.L. and Gerba, C.P. (2004) Tracking the concentration of heterotrophic plate count bacteria from the source to the consumer's tap. International Journal of Food Microbiology 92(3), 289-295.

Pinto, A., Xi, C. and Raskin, L. (2012) Bacterial community structure in the drinking water microbiome is governed by filtration processes. Environmental Science & Technology.

Poitelon, J.B., Joyeux, M., Welte, B., Duguet, J.P., Prestel, E., Lespinet, O. and Dubow, M.S. (2009) Assessment of phylogenetic diversity of bacterial microflora in drinking water using serial analysis of ribosomal sequence tags. Water Research 43(17), 4197-4206.

Poitelon, J.B., Joyeux, M., Welte, 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 & Biotechnology 37(2), 117-128.

Polaczyk, A.L., Narayanan, J., Cromeans, T.L., Hahn, D., Roberts, J.M., Amburgey, J.E. and Hill, V.R. (2008) Ultrafiltration-based techniques for rapid and simultaneous concentration of multiple microbe classes from 100-L tap water samples. Journal of Microbiological Methods 73(2), 92-99.

Reasoner, D.J. and Geldreich, E.E. (1985) A new medium for the enumeration and subculture of bacteria from potable water. Applied and Environmental Microbiology 49(1), 1-7.

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., Stres, B., Thallinger, G.G., Van Horn, D.J. and Weber, C.F. (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology 75(23), 7537-7541.

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.

Sheu, S.-Y., Jiang, S.-R., Chen, C.A., Wang, J.-T. and Chen, W.-M. (2011) Paracoccus stylophorae sp. nov., isolated from the reef-building coral Stylophora pistillata. International Journal of Systematic and Evolutionary Microbiology 61(9), 2221-2226.

Simoes, L.C., Simoes, M. and Vieira, M.J. (2010) Adhesion and biofilm formation on polystyrene by drinking water-isolated bacteria. Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology 98(3), 317-329.

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.

Stewart, M.H., Wolfe, R.L. and Means, E.G. (1990) Assessment of the bacteriological activity associated with granular activated carbon treatment of drinking water. Applied and Environmental Microbiology 56(12), 3822-3829.

Sun, L.-N., Zhang, J., Kwon, S.-W., He, J., Zhou, S.-G. and Li, S.-P. (2012) Paracoccus huijuniae sp. nov., an amide pesticides-degrading bacterium isolated from activated sludge of a wastewater bio-treating system. International Journal of Systematic and Evolutionary Microbiology.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution 28(10), 2731-2739.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25(24), 4876-4882.

Turner, S., Pryer, K.M., Miao, V.P.W. and Palmer, J.D. (1999) Investigating deep phylogenetic relationships among cyanobacteria and plastids by small submit rRNA sequence analysis. Journal of Eukaryotic Microbiology 46(4), 327-338.

Wang, Q., Garrity, G.M., Tiedje, J.M. and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73(16), 5261-5267.

Weon, H.-Y., Yoo, S.-H., Kim, S.-J., Kim, Y.-S., Anandham, R. and Kwon, S.-W. (2010) Massilia jejuensis sp. nov. and Naxibacter suwonensis sp. nov., isolated from air samples. International Journal of Systematic and Evolutionary Microbiology 60(8), 1938-1943.

White, C., Tancos, M. and Lytle, D.A. (2011) Microbial Community Profile of a Lead Service Line Removed from a Drinking Water Distribution System. Applied and Environmental Microbiology 77(15), 5557-5561.

Williams, M.M., Domingo, J.W.S., Meckes, M.C., Kelty, C.A. and Rochon, H.S. (2004) Phylogenetic diversity of drinking water bacteria in a distribution system simulator. Journal of Applied Microbiology 96(5), 954-964.

Xu, P., Li, W.-J., Tang, S.-K., Zhang, Y.-Q., Chen, G.-Z., Chen, H.-H., Xu, L.-H. and Jiang, C.-L. (2005) Naxibacter alkalitolerans gen. nov., sp. nov., a novel member of the family 'Oxalobacteraceae' isolated from China. International Journal of Systematic and Evolutionary Microbiology 55(3), 1149-1153.

Zhang, Y., Griffin, A. and Edwards, M. (2008) Nitrification in premise plumbing: Role of phosphate, pH and pipe corrosion. Environmental Science & Technology 42(12), 4280-4284.

Zwart, G., Crump, B.C., Agterveld, M., Hagen, F. and Han, S.K. (2002) Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. Aquatic Microbial Ecology 28(2), 141-155.

Chapter 5. Influence of Source Water, Filtration Technology, and Disinfection on Drinking Water Bacterial Community

5.1. Abstract

To investigate the influence of source water, disinfection and filtration technology on tap water bacterial community, four drinking water treatment plants were sampled from the raw river to customer's tap. The membrane filtration and conventional sand filtration processes were compared as to their bacterial community changes in plant scale. Bacterial community dynamics during drinking water treatment and distribution processes were investigated through 16S rDNA based pyrosequencing analysis. Membrane filtration had better treatment efficiency than sand filtration, however, didn't cause significantly different bacterial communities. Chlorination was the key step controlling the bacterial community structure in tap water. The influence order on the water bacterial community were disinfection > water sources > treatment techniques. Sometimes the influence of distribution system was greater than treatment processes and water sources. *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, and *Firmicutes* were found dominated in all the water treatment plants. The persistent core bacteria population which survived each treatment and distribution steps and appeared in all the four water treatment plants was from the genus of *Sphingomonas*. The core bacterial populations observed in all the finished water and tap water samples were belonged to the family of *Sphingomonadaceae*, *Comamonadaceae*, *Moraxellaceae*, and *Pseudomonadaceae*, which were classified to the bacteria genera of *Sphingomonas*, *Acinetobacter*, and *Pseudomonas*.

5.2. Introduction

Microbial contamination in drinking water causes pipe corrosion, water quality deterioration and outbreak of waterborne disease (Craun et al. 2010, Li et al. 2010, White et al. 2011). Previous surveys showed that from the year 1971 to 2006, 33% waterborne diseases were cause by untreated source water, 39% by inadequate or interrupted treatment process and 18% by distribution system and premise plumbing deficiencies (Craun et al. 2010). Source water was considered the original seed for tap water microbial community since the most tap water bacteria were found to be fresh water origin (Henne et al. 2012, Poitelon et al. 2009). Therefore the primary objective for drinking water treatment is to remove microorganisms in water. Some research showed that tap water community was shaped by filtration process (Pinto et al. 2012). Some found disinfection step governed tap water bacterial community (Eichler et al. 2006, Poitelon et al. 2010). Others suggested that water distribution system had significant effects on drinking water microbial community (Lautenschlager et al. 2010). There is no consistent answer for the key factor that controls the drinking water bacterial community.

The application of microfiltration (MF) and ultrafiltration (UF) was rapidly increased in recent years due to the good reputation for particles and microorganisms removal as well as disinfection by-products formation reduce by reducing the required disinfection dose (Alspach et al. 2008, Jacangelo and Watson 2002). The physical size of protozoan cysts and oocysts, bacteria and virus are in the range of 1 ~ 15 μm, 0.5 ~ 10 μm and 0.02 ~ 0.08 μm, respectively. MF and UF are designed to be able to reject particles large than the membrane pore size (are 0.05 ~ 5 μm for MF and 0.005 ~ 0.05 μm for UF). Many reports showed that MF and UF achieved up to 7-log removal for particles and pathogens (Jacangelo and Watson 2002). Some results suggested that membrane filtration had better efficiency for microorganism removal than conventional

filtration process (Ho et al. 2012) . However the defects in membrane fiber may allow microbes pass through the sieving barrier. The attached microorganisms may cause fouling problem (Guo et al. 2010). The comparisons on whole microbial community composition changes before and after filtration were still very rare. The plant scale comparisons with conventional filtration haven't been reported.

In this study we selected four water treatment plants including two membrane filtration plants and two conventional filtration plants taking two different rivers as source water. So that in one river we have two different parallel treatment streams with one conventional filtration and one membrane filtration for comparison. Water samples were taken after different treatment steps from the raw water to customer's tap water. The microbial communities were investigated through 16S rRNA gene based pyrosequencing analysis. The aim of this study are to investigate the microbial community dynamics during water treatment process; find the major factor that finally controlled the tap water bacterial community; find the core bacterial community that can survive the treatment process and persistent in drinking water.

5.3. Material and Methods

5.3.1. Water sampling in two drinking water treatment plants

Four drinking water treatment plants (C, D, E and F) located in Tennessee were sampled on 2012, May 14, May 21, June 18 and June 25, respectively. Water treatment plant C and D draw their raw water from the same river. The raw water treated in plant E and F are from the same source river. The raw water was treated through pre-disinfection, coagulation, flocculation, filtration, and chlorination. Plant C didn't have pre-disinfection. Plant D used 1.4 mg/L chlorine

for the raw water pre-chlorination. Plant E and F sequentially added 0.3 ~ 0.5 mg/L sodium permanganate and chlorine dioxide and chlorine to disinfect the raw water. The coagulant agent used in plant C and E were aluminum chlorohydrate. Plant D added polyaluminium chloride for coagulation. Plant F selected aluminum chloride hydroxide sulfate and polyaluminium hydrochlorosulfate as coagulants. Water treatment C and E are membrane filtration plants using GE ZeeWeed hollow-fiber ultrafiltration membrane system and submerged microfiltration Siemens Memcor® membranes for the filtration process, respectively. The membrane filters were cleaned use air scouring (once per min) and back plusing every 20 min in plant C and every 45 min in plant E. Both membrane filtration plants use 25 mg/L chlorine and citric acid (pH 2.0) for the periodical membranes Cleaning-in-Place (CIP) every month or as needed according to the permeate pressure in some situations. Water treatment D and F are conventional sand filtration plants and stay with the conventional multiple-layer sand filters for their filtration processes. Plant D had 4 filters and back washed one sand filters per night. Plant F had 10 filters and back washed 5 in one day, one wash per week. All of four treatment plants use chlorine for their disinfection processes. Water samples taken from source river to customer's tap along treatment and distribution processes were raw water pumped into water treatment plant, water after filtration (either after membrane or sand filters), water after chlorination, tap water after distribution. The fresh tap water after distribution was taken from customer's tap faucet after flushing the faucet for 10 minutes at maximum flow. Two liters of raw water and 100 liters of water after each treatment steps and tap water were collected in autoclaved carboys for biological analysis. One liter of water from each sample was used for water quality analysis.

5.3.2. Water quality analysis

Water quality parameters during water treatment processes were measured including pH, turbidity, conductivity, free chlorine, dissolved organic carbon (DOC), sulfate, nitrate, and chloride. Turbidity and conductivity were measured with Hach 2100N turbidimeter (Hach Company, Loveland, Colorado, USA) and Orion model 122 conductivity meters (Orion Research Inc., Boston, Massachusetts, USA), respectively. Free chlorine was quantified using the standard "4500-Cl F" DPD Ferrous Titrimetric Method (APHA 2005). Dissolved organic carbon was analyzed using the Shimadzu SSM-5000A TOC analyzer (Shimadzu Corporation, Kyoto, Japan) as described in the standard method.

5.3.3. Bacteria collection and DNA extraction

Bacteria were harvested within less than 4 hours of sampling by filtration of 2 liters of raw water on a 0.22 µm pore size polycarbonate filters and 100 liters of treated water with a tangential-flow ultrafiltration system as previously described. All the tubes and containers included in the ultrafiltration system were disinfected with 10% hypochlorous acid, washed with deionized water, and autoclaved before use. Chlorinated water samples were dechlorinated by adding sodium thiosulfate to a final concentration of 50 mg/L and sodium polyphosphate was added to each water sample to a final concentration of 0.01% (w/v) as the dispersant as previously described (Hill et al. 2007, Mull and Hill 2009, Polaczyk et al. 2008). Bacteria were collected on a 0.22 µm pore size polycarbonate filters by filtration of the ultrafiltrated water. The filter sandwiches were stored at -80 °C for further analysis. The whole genome DNA was extracted from the filter sandwiches using FastDNA spin kit for soil (MP Biomedicals, Santa Anna, CA).

5.3.4. Pyrosequencing of 16S rDNA amplicons

16S rRNA gene amplicon libraries were generated for bacterial samples collected after each treatment steps with primers targeting the V4 hypervariable region: 515F (5'-GTGCCAGCAGCCGCGGTAA-3') and 806R (5'-GGACTACCAGGGTATCTAAT-3') (Nikkari et al. 2002). The forward and reverse primer included a Roche 454 A and B pyrosequencing adapter (454 Life Sciences, Branford, CT, USA). A 10-bp barcode sequence which is unique to each individual sample linked in the forward side. Polymerase Chain Reaction (PCR) amplification was performed in a volume of 50 μL reaction system, each containing 5 μL of FastStart High Fidelity Reaction Buffer with 1.8 mM $MgCl_2$, 200μM dNTPs, 0.4 μM forward and reverse primers, 10-100 ng of DNA, and 2.5 U FastStart High Fidelity Enzyme Blend (Roche Diagnostics, Germany). The thermal cycling program used for PCR was: 94 °C for 3 min for 1 cycle; 94 °C for 30 sec, 57 °C for 1 min, 72 °C for 2 min for 20 cycles; and a final extension at 72 °C for 2 min. PCR products were purified with the Qiagen PCR purification kit (Qiagen, Valencia, California, USA) and the Agencourt AMPure PCR purification system (Beckman Coulter, Danvers, Massachusetts, USA). The quality of each amplicon library was measured using the Agilent DNA 7500 kit with a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, USA). Equal molar of amplicons from each sample were pooled together, and then amplified through emulsion PCR using the GS FLX titanium emPCR amplicon kit according to the manufacturer's protocols (454 Life Sciences, Branford, CT, USA). PCR products were sequences at the Center for Environmental Biotechnology at the University of Tennessee using a 454 Genome Sequencer FLX titanium platform (454 Life Sciences, Branford, CT, USA).

5.3.5. Data processing and statistical analysis

The analysis of 16S rRNA gene sequences obtained from pyrosequencing were performed with MOTHUR v.1.27.0 (Schloss et al. 2009). Sequences were denoised, sorted and trimmed according to sample-specific barcodes. After dereplication, low quality, short (<100 bp) and chimera sequences were removed following MOTHUR pipeline as described previously (Huse et al. 2010). The non-redundant sequences were aligned against the SILVA database. Operational taxonomic units (OTUs) were assigned with the average neighbor clustering algorithm and taxonomy was assigned using RDP database with a bootstrap cutoff of 80% (Wang et al. 2007). Hierarchical cluster analysis and principle coordinate analysis (PCoA) was performed using weighted UniFrac metrices. Detrended correspondence analysis (DCA) was performed to investigate the bacteria community variations among different water samples. DCA for bacterial family variables was performed by using CANOCO 4.5 (Microcomputer Power, Ithaca, NY). A relative abundance above 2% at family level were used to reduce the weight of rare species,

Nucleotide sequence accession numbers

454 pyrosequencing raw data was submitted to NCBI GenBank Short Read Archive (SRA) with the accession number SRA058624.

5.4. Results and Discussion

5.4.1. Water quality changes after each treatment steps

The water quality changes during treatment, distribution, and stagnation processes were shown in Table 5.1. The raw water in water treatment plant E and F had relatively higher turbidity, DOC

and conductivity than the raw water in plant C and D. Filtration step greatly improved water turbidity. The filtered water had the lowest turbidity. The membrane permeate water samples had lower turbidity than the sand filtration samples. After chlorination and distribution, water turbidity gradually increased. The tap water from different water treatment plants had similar turbidity. The pH values were controlled between 7.0 -8.0. The free chlorine residual in the finished water were controlled between 2.20 - 2.50 mg/L. After distribution the free chlorine always decreased but still maintained at 1.94 - 2.24 mg/L. In summary the physical and chemical water quality controlled in finished water and tap water from all the four plants were very close to each other. According to the HPC number, membrane filtration had better bacterial removal than conventional sand filtration. Combined with coagulation membrane filtration in plant C achieved 3.0-log HPC removal, while after sand filtration plant D had 1.4-log HPC removal for the treatment of the same source river water. After membrane filtration plant E achieved 6.6-log HPC removal, while plant F after sand filtration obtained 5.5-log HPC removal. The predisinfection with strong oxidants in plant E and F may contribute to the higher HPC removal efficiency than plant C and D. The primary disinfection in plant C and D achieved another 3.9 and 4.4-log HPC removal. However the last step of chlorination in plant E and F didn’t cause significant change for HPC number.

Table 5.1 Summary of water quality data during treatment and distribution processes

	Temperature (°C)	pH	Turbidity (NTU)	Free Cl_2 (mg/L)	DOC (mg/L)	Conductivity (uS/cm)	HPC (CFU/100 ml)
C_raw	17.5	7.74	4.68	-	1.63	127	$2.27\times10^7 \pm 9.90\times10^5$
C_membrane	17.5	7.98	0.016	-	1.04	128	$2.50\times10^4 \pm 1.41\times10^3$
C_finish	17.5	7.95	0.016	2.25	1.22	153	3.2 ± 1.5
C_tap	22.8	7.90	0.044	1.94	1.20	147	85.3 ± 11.7
D_raw	19.5	7.38	1.50	-	1.00	101	$1.18\times10^7 \pm 2.83\times10^5$
D_sand	20.1	7.08	0.027	-	0.84	105	$5.07\times10^5 \pm 1.61\times10^5$
D_finish	22.0	7.07	0.034	2.30	1.32	106	21.0 ± 4.2
D_tap	19.6	7.02	0.045	2.09	1.19	117	11.0 ± 3.1
E_raw	27.4	7.62	53.9	-	2.29	320	$8.10\times10^6 \pm 4.24\times10^5$
E_membrane	28.1	7.21	0.015	-	2.83	324	2.0 ± 0.0
E_finish	27.2	7.30	0.040	2.20	2.14	330	101.2 ± 0.0
E_tap	26.0	7.29	0.046	1.95	2.35	340	12.6 ± 2.5
F_raw	21.0	7.80	7.00	-	1.50	298	$3.15\times10^6 \pm 7.78\times10^5$
F_sand	22.4	7.77	0.024	-	2.01	316	9.7 ± 4.7
F_finish	23.0	7.70	0.050	2.50	1.86	314	29.7 ± 4.7
F_tap	25.4	7.66	0.050	2.24	1.62	311	24.0 ± 8.5

5.4.2. Overview of bacterial community composition in drinking water treatment and supply systems

Bacterial community changes during treatment and distribution process were analyzed through pyrosequencing of the 16S rRNA gene amplicons. Total of 13359, 17507, 31587 and 27704 sequences, with an average length of ~250 bp, were obtained from water treatment plant C, D, E and F respectively. At 80% cut off value, these sequences were assigned to 14, 17, 19 and 18 bacterial in water treatment plant C, D, E and F, respectively. Bacterial members in *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, and *Firmicutes* were found dominated in all the water treatment plants. *Proteobacteria* was the most dominant phylum accounting for 46.9% ~ 74.7% total sequences in each the water treatment plants. *Bacteroidetes* and *Actinobacteria* were the second and third largest phylum accounting for 9.8% ~ 34.4% and 7.5% ~ 26.1% total sequences for each treatment plants.

The relative abundance of bacterial phyla and bacterial class in *Proteobacteria* at each sampling location showed in Figure 5.1 A and B. Bacterial community structures were greatly changed after each treatment step. Filtered water samples with the same sources were close to each other. All the chlorinated water showed similar patterns. The raw water communities especially those samples from the same river were similar to each other. Raw water in plant C and D were dominated by *Bacteroidetes* (40.6% and 48.6%), and *Betaproteobacteria* (29.1% and 38.0%). Bacteria members from *Actinobacteria* (33.7% and 31.5%), *Betaproteobacteria* (22.4% and 20.7%), and *Bacteroidetes* (19.1% and 22.7%) were the dominant populations in the raw water of plant E and F. After filtration *Alphaproteobacteria* (78.1% and 27.8%) and *Actinobacteria* (14.9% and 27.6%) became the dominant population in plant C and D. *Betaproteobacteria* (53.8% and 48.6%) and *Gammaproteobacteria* (23.7% and 18.1 %) were dominated in plant E and F. *Betaproteobacteria* (11.0% ~ 32.6%), *Gammaproteobacteria* (3.9% ~ 44.0%), *Actinobacteria* (9.4% ~ 27.0%) and *Firmicutes* (0.8% ~ 25.8%) were found dominated in finished water. *Betaproteobacteria* (0% ~ 30.9%), *Gammaproteobacteria* (16.3% ~ 99.9%), and *Actinobacteria* (0% ~ 56.1%) were observed with high relative abundance in tap water. During treatment and distribution processes the relative abundance of *Bacteroidetes* decreased and *Gammaproteobacteria* increased.

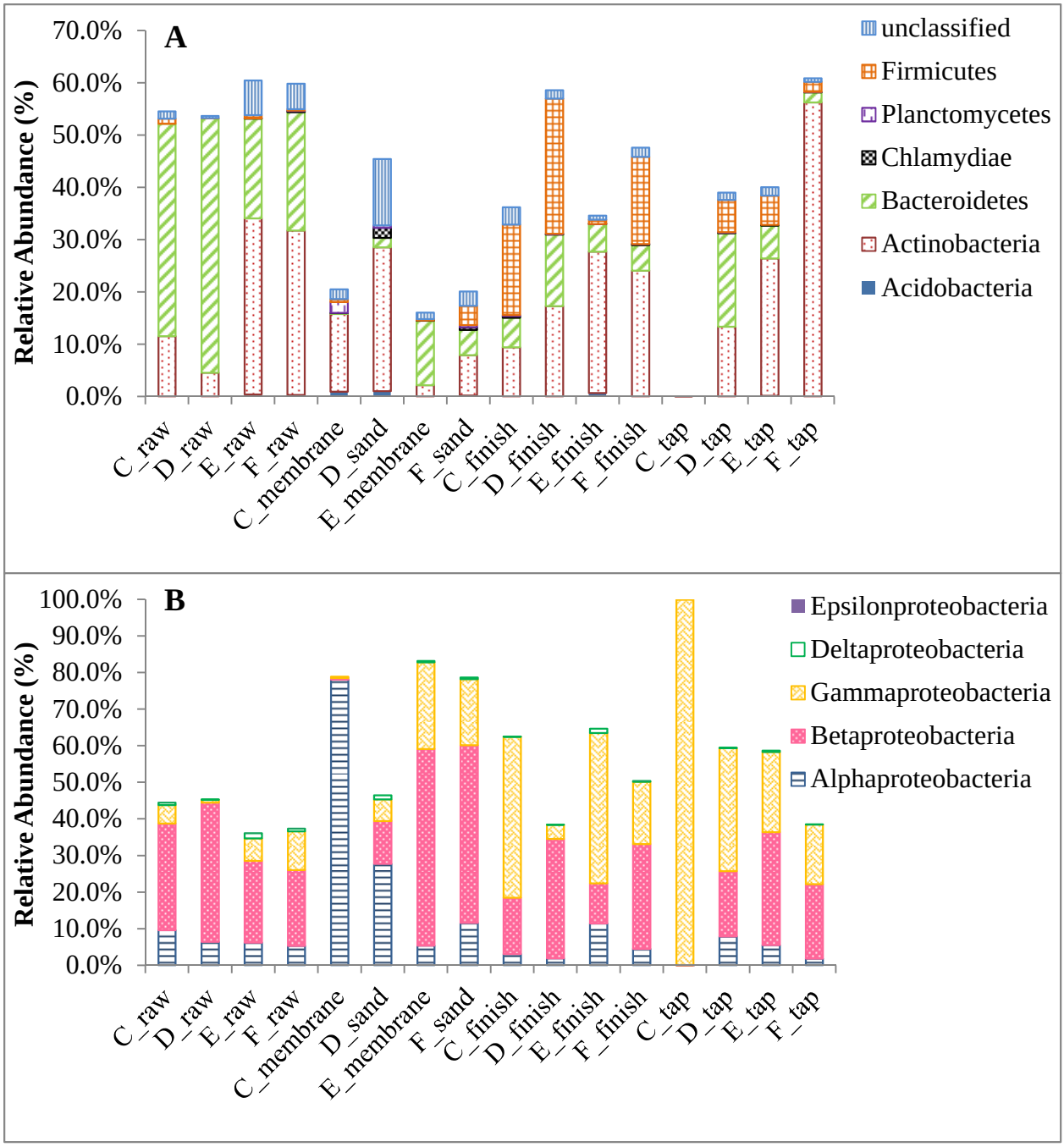

Figure 5.1 Bacterial community composition identified by 16S rRNA gene pyrosequencing. Bacteria phyla (A) and *Proteobacteria* classes (B) with relative abundance > 1% are shown. The raw water in treatment plant C and D are from the same source river. The raw water in treatment plant E and F are from the same source river. Water treatment plant C and E are membrane filtration plants. Water treatment plant D and F are conventional sand filtration plants.

5.4.3. Comparison of bacterial population changes in membrane filtration processes with conventional sand filtration processes

To compare the treatment differences caused by different treatment techniques, two water treatment plants use the same source river were put together for further analysis. Sequences were further classified into family and genus level. Differences and connections between microbial communities after different treatment steps were analyzed through DCA analysis. Bacterial community dynamics were illustrated in DCA plot based on bacterial family finger prints (Figure 5.2 A and B). Only abundant bacterial families (>2%) were included in the DCA analysis with 59.2% ~ 99.9% sequences in each water samples were covered. In both figures bacterial community from raw to filtration and disinfection were moved to the same direction. From finished water to tap water bacterial community shifted to a different direction. The raw water from same river had similar community composition and always clustered together. No obvious differences were observed between membrane filtration and sand filtration as to their bacterial communities since they were close to each other in both cases. The finished water and tap water were separated from the raw and filtered water.

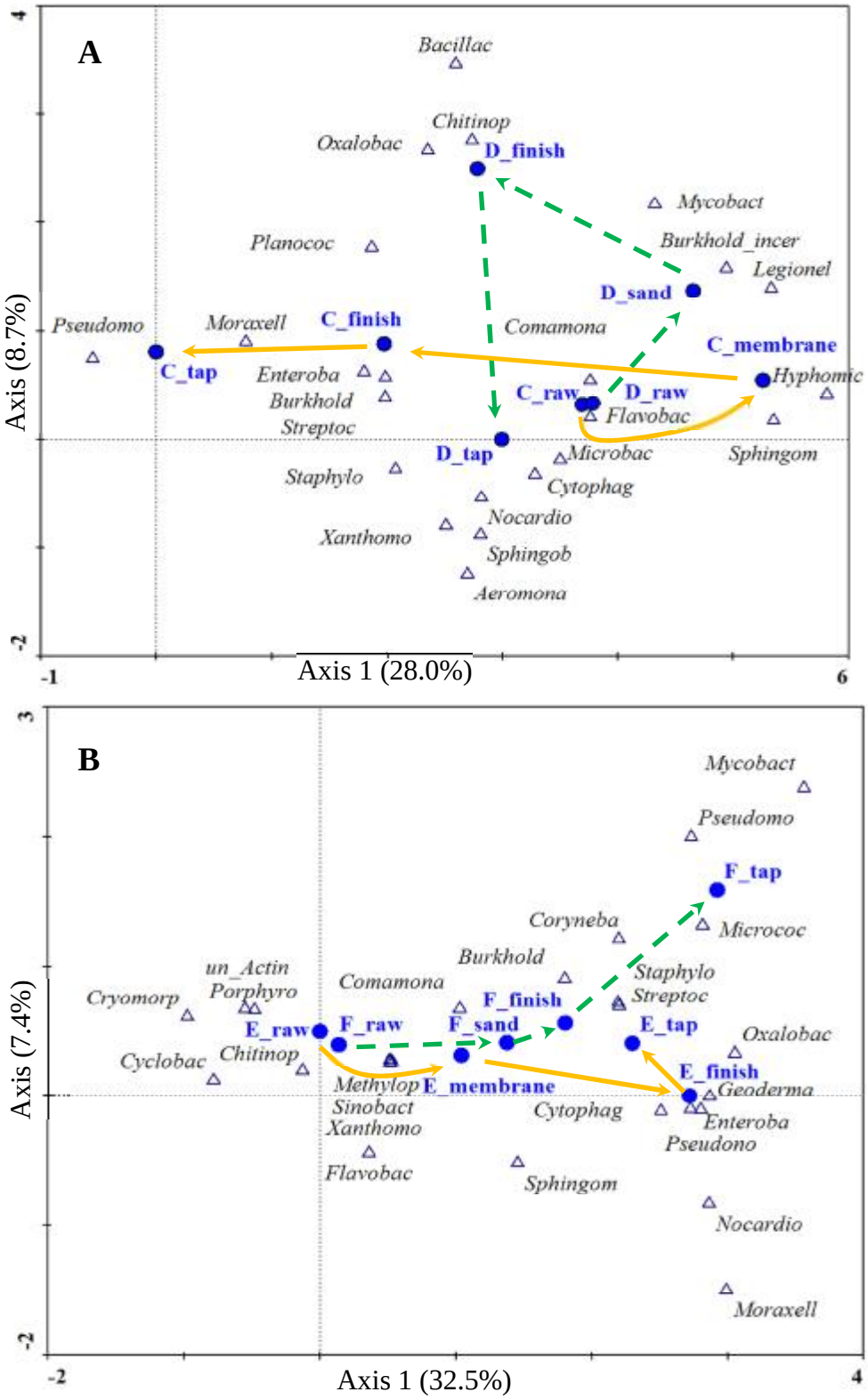

Figure 5.2 Detrended correspondence analysis (DCA) for abundant bacterial families (>2%) in plant C and D (A) and E and F (B). The raw water in treatment plant C and D are from the same source river. The raw water in treatment plant E and F are from the same source river. Water

treatment plant C and E are membrane filtration plants. Water treatment plant D and F are conventional sand filtration plants.

Axis 1 in Figure 5.2 A reflected the bacterial community shift after filtration, chlorination, and distribution of plant C. Bacterial community change after filtration and chlorination in plant D followed a similar direction but reflected by both Axes. The shift from finished water to tap caused by distribution was mainly along Axis 2. The raw water in plant C and D were very close to each other because they contain similar bacteria families. As shown in Figure 5.2 A the dominant bacterial families were *Flavobacteriaceae* (30.7% and 38.9%), *Comamonadaceae* (17.5% and 29.6%), *Microbacteriaceae* (7.7% and 2.8%) and *Cytophagaceae* (2.4% and 6.2%), which appeared near the raw water sample dots. *Flavobacteriaceae* was classified into the genus of *Flavobacterium* (30.1% and 38.9%). Water after membrane filtration and sand filtration were also similar to each other reflected as their close position in the DCA plot. They were clustered because they both have high abundance of *Mycobacteriaceae* (14.2% and 26.3%) and *Sphingomonadaceae* (55.9% and 17.7%) and low abundance of *Legionellaceae* (0.02% and 2.2%). They were separated because the difference on the relative abundance of *Hyphomicrobiaceae* (19.4% and 3.4%) and *Burkholderiales*_incertae_sedis (0.1% and 6.1%). These bacterial families were classified into the genus of *Mycobacterium* (14.2% and 26.3%), *Sphingomonas* (51.9% and 15.2%), *Legionella* (0.02% and 2.2%), *Hyphomicrobium* (19.3% and 3.3%), and *Aquabacterium* (0.0% and 5.8%). The two finished water separated along Axis 2 which only account for 8.7% of total variances. The separation was caused by the differences in the relative abundances of their dominant populations. The dominant bacterial in the finished water in plant C were *Enterobacteriaceae* (25.2%), *Moraxellaceae* (13.7%) and *Oxalobacteraceae* (7.8%), while the dominant bacterial families of finished water in plant D

were *Oxalobacteraceae* (28.9%), *Bacillaceae* (19.8%), *Mycobacteriaceae* (13.4%) and *Chitinophagaceae* (12.0%). Those families can be found in the finished water but appeared with different abundances. The corresponding genera in finished water of plant C were *Escherichia_Shigella* (24.4%), *Acinetobacter* (13.0%), and *Naxibacter* (7.4%). The dominant genera in finished water of plant D were identified as *Naxibacter* (25.8%), *Mycobacterium* (13.4%), and *Sediminibacterium* (12.0%). The tap water generated from plant C was composed of *Pseudomonadaceae* (58.7%) and *Moraxellaceae* (41.1%). The tap water from plant D was composed of *Aeromonadaceae* (19.0%), *Xanthomonadaceae* (10.8%) and many bacterial families in the raw water. *Moraxellaceae, Aeromonadaceae,* and *Xanthomonadaceae* were classified into the genus of *Acinetobacter* (41.0%), *Aeromonas* (19.0%), and *Stenotrophomonas* (7.1%), respectively. Therefore the tap water from plant D was more close to the raw water and far from the tap water from plant C. Both tap water were far away from their upper source, the corresponding finished water, indicating the substantial influence caused by water distribution system. Eichler also found one of the tap water was similar to the raw water in his survey by sampling along water treatment processes (Eichler et al. 2006). This probably caused by the sloughing biofilm inside water distribution systems. Biofilm attached on the water pipes were considered a bacteria reservoir. The detached biofilm may get into the bulk water and finally transported to customer's tap through the distributed water flow. Biofilm usually has high cell density compared to bulk water, consequently will have substantial effects on the bulk water bacterial community. Biofilm bacteria were originally seeded by the raw river water. Therefore when we only consider the first several dominated population the tap water may be more close to the raw water.

The bacteria community shifts in plant E and F were consistent with each other. From the raw water to the filtered water and the finished water, they both moved along Axis 1 (Figure 5.2 B). The tap water separated from other samples along both axes. Similar to what we observed before, the two raw water were similar and the filtered water were also close to each other, suggesting that membrane filtration and sand filtration didn't cause great changes in bacterial communities. The first two dominant populations in the two rivers were unclassified *Actinomycetales* (30.1% and 25.9%) and *Comamonadaceae* (5.9% and 12.3%). Other bacterial families on the left side of the two raw water samples in the DCA plot were mostly in relative low abundance and observed in other water samples. The two filtered water were close to each other because they share the predominant family *Comamonadaceae* (26.9% and 40.9%), which was not classifiable at genus level in plant E but identified as *Curvibacter* (27.6%) in plant F. High abundance of *Methylophilaceae* (16.1%) and *Sinobacteraceae* (11.8%) were observed in the permeate water in plant E, which were classified into the genera of *Methylophilus* (16.0%) and *Hydrocarboniphaga* (11.8%), respectively. *Sphingomonadaceae* (10.1%) and *Enterobacteriaceae* (8.6%) appeared in sand filtered water with high abundance, which were classified into the genera of *Sphingomonas* (8.4%) and *Escherichia_Shigella* (8.2%), respectively. *Moraxellaceae* (29.4%), *Mycobacteriaceae* (18.2%), *Sphingomonadaceae* (10.3%) and *Oxalobacteraceae* (8.6%) were found dominant in the finished water in plant E, which were classified into the genera of *Acinetobacter* (29.1%), *Mycobacterium* (18.1%), *Sphingomonas* (9.9%), and *Massilia* (4.2%), respectively. *Comamonadaceae* (19.0%), *Staphylococcaceae* (9.0%) and *Corynebacteriaceae* (7.7%) were found dominant in the finished water in plant F, which were classified into the genera of *Curvibacter* (11.9%), *Staphylococcus* (9.0%), and *Corynebacterium* (7.7%), respectively. *Mycobacteriaceae* (8.2% and 47.5%) and

Comamonadaceae (12.6% and 6.2%) were observed in the tap water from plant E and F, which were identified as *Mycobacterium* (8.1% and 47.5%) and *Curvibacter* (7.7% and 4.0%). *Moraxellaceae* accounted for 10.1% in the tap water from plant E, was assigned to the genus of *Acinetobacter* 8.3%. The tap water in plant F also had high abundance of *Oxalobacteraceae* (8.6%) and *Pseudomonadaceae* (9.4%), only small amount of which were classified into the genus of *Massilia* (0.4%) and *Pseudomonas* (0.4%).

Some research showed that membrane filtration process had better bacterial removal efficiency and cause different bacteria community structure in the finished water (Bottino et al. 2001, Ho et al. 2012). In our study membrane filtration process and sand filtration process didn't cause great differences in bacterial communities in the filtrated water. Lager differences were observed after chlorination and distribution, suggesting that chlorination and distribution had greater influence than filtration. However, previous reports were from pilot scale studies. In pilot scale the preconditions can be controlled precisely. In reality it's difficult to compare the two treatment processes on the same basis because no water treatment plants use two different treatment streams at the same time. In plant scale operation, many factors could affect the treatment efficiency and bacterial community in the finished water, such as raw water pretreatment, the defects on membrane and fouling issues etc. (Peter-Varbanets et al. 2011, Walsh et al. 2009, Wang et al. 2011). Our study is the first time compare membrane filtration treatment process with conventional sand filtration process as to the bacterial community structures under very close preconditions. In this study, only four plants were sampled at short intervals. More repetitive works need to be performed and more water quality parameters need to be considered for the evaluation of the performance of membrane filtration versus conventional filtration processes. Several pathogens were detected in many water samples. However further

research is needed to elucidate their activity and functions. The 16S rDNA based analyses may overestimate the potential risks due to the fact that both live and dead cells were included at DNA level detection and identification.

5.4.4. Persistent and core populations detected during drinking water treatment and distribution processes

Source water was considered the original seed from tap water bacteria. Some bacteria populations were able to survive each treatment steps and be persistent from the raw to the tap water. Taxonomic analysis revealed a broad range of persistent bacteria. There are 4, 18, 37, and 55 persistent bacterial families were detected in plant C, D, E and F, respectively. There are 15 persistent bacterial families detected in at least 3 plants with 4 observed in all the surveyed plants. Total 9 out of 15 were identified as known bacterial families associated with *Nocardioidaceae* (0.03% ~ 2.6%), *Chitinophagaceae* (0.1% ~ 12.0%), *Caulobacteraceae* (0.02% ~ 0.3%), *Hyphomicrobiaceae* (0.01% ~ 3.4%), *Burkholderiales*_incertae_sedis (0.02% ~ 6.1%), *Oxalobacteraceae* (0.3% ~ 28.9%), *Pseudomonadaceae* (0.02% ~ 9.4%), *Sphingomonadaceae* (0.03% ~ 55.9%), and *Comamonadaceae* (0.03% ~ 40.9%). The last two bacteria families appeared in all the four plants. The other two shared by all the plants belonged to the order of *Actinomycetales* (0.03% ~ 30.1%) and the class of *Gammaproteobacteria* (0.03% ~ 3.4%) which were not close to any known bacterial families. Genus level classification identified 3, 18, 34 and 53 persistent bacterial genera in plant C, D, E and F, respectively. Total 12 persistent core bacterial genera appeared in at least 3 plants with 3 observed in all the surveyed plants. Only 3 out 12 genera were close to known bacteria genus of *Nocardioides* (0.02% ~ 2.2%), *Brevundimonas* (0.02% ~ 0.2%), and *Sphingomonas* (0.03% ~ 51.9%). Bacterial from the genus of *Sphingomonas* was the only persistent core population that was detected from source to the

tap and widely spread in all the four plants.

The bacterial population persisted in finished water and tap water had potential risks for water quality and human health. Total 104 bacterial families and 147 genera were detected in the finished water from four plants. The core populations shared by finished water generated from all the water treatment plants accounted for 34.6% and 20.4% of total taxonomic units at family level and genus level, respectively. Only 12 out of 30 core bacterial populations were assigned to known bacterial genera: *Mycobacterium* (0.2% ~ 18.1%), *Nocardioides* (0.1% ~ 2.2%), *Streptomyces* (0.02% ~ 0.1%), *Flavobacterium* (0.1% ~ 2.0%), *Brevundimonas* (0.04% ~ 0.1%), *Sphingomonas* (0.7% ~ 9.9%), *Ralstonia* (0.1% ~ 6.3%), *Pelomonas* (0.04% ~ 1.1%), *Acinetobacter* (0.7% ~ 29.1%), *Pseudomonas* (0.2% ~ 1.7%), *Staphylococcus* (0.04% ~ 9.0%) and *Turicibacter* (0.03% ~ 0.3%). Total 100 bacterial families and 142 genera were detected in the tap water from four plants. Only a small amount of core populations (6% at family level and 4.2% at genus level) were found in the tap water samples. The core populations observed in tap water were all covered by the core populations in finished water. Total 6 core bacterial families were observed in all the finished water and tap water samples: *Sphingomonadaceae* (0.03% ~ 10.3%), *Comamonadaceae* (0.03% ~ 19.0%), *Moraxellaceae* (0.5% ~ 41.1%), *Pseudomonadaceae* (0.2% ~ 58.7%), unclassified *Gammaproteobacteria* (0.03% ~ 0.6%), and unclassified *Actinomycetales* (0.03% ~ 2.6%). Total 6 core bacteria genera were observed in all the finished and tap water samples which belonged to 3 know bacteria genera: *Sphingomonas* (0.03% ~ 9.9%), *Acinetobacter* (0.5% ~ 41.0%), and *Pseudomonas* (0.2% ~ 3.0%).

5.4.5. Chlorination controlled the bacterial community structure in tap water

Many factors could affect the bacterial community structure, such as treatment processes, source water, different filtration technology, and distribution systems. In order to find the main factor

that finally controlled the tap water bacterial communities, all the water samples were put together for further analysis, including all abundant as well as rare species. Hierarchical cluster analysis and PCoA analysis were performed based on weighted and unweighted the UniFrac distance generated from the whole database. In weighted cluster analyses, most finished water and tap water were clustered together and separated from the raw water and filtrated water, suggesting that chlorination had substantial effect on the bacterial community structure (Figure 5.3 A). Inside each branch, water samples clustered firstly according to treatment steps and then according to their sources instead of treatment method, suggesting that the influences order on the water bacterial community were treatment processes larger than water sources and larger than treatment techniques did. However the tap water samples were not clustered with the corresponding finished water and sometimes were far away from the disinfection branch, suggesting that the influence of distribution system was greater than source water, sometimes even greater than treatment processes.

In the unweighted cluster tree, the major shift of bacterial community was also caused by chlorination, as most chlorinated water were close to each other and the raw water and some filtered water were separated to the chlorination branch (Figure 5.4 A). The unweighted tree was also clustered firstly according to treatment steps and then according to their sources. The membrane filtered water samples were not clustered neither did the sand filtered water samples. Source water had greater effects on the unweighted tree than the weighted tree. The tap water samples tend to group by source suggesting the seed of the tap water bacteria can be traced from water treatment plants. However the tap water samples were not always clustered with their corresponding finished water samples and the chlorination branch, suggesting that the influence of distribution system was greater than source water and treatment processes.

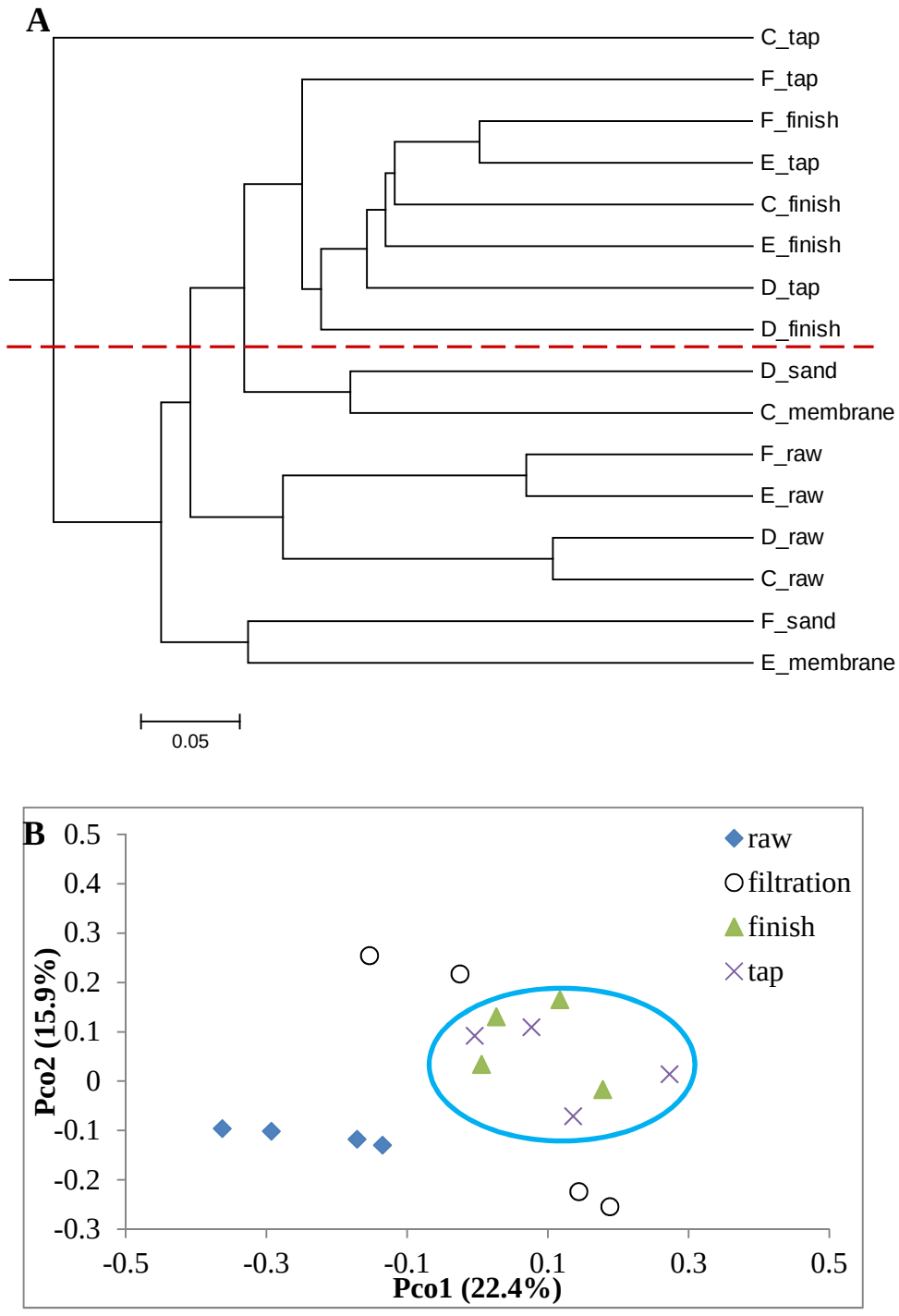

Figure 5.3 Hierarchical cluster analysis (A) and PCoA analysis (B) based on weighted UniFrac distance matrix.

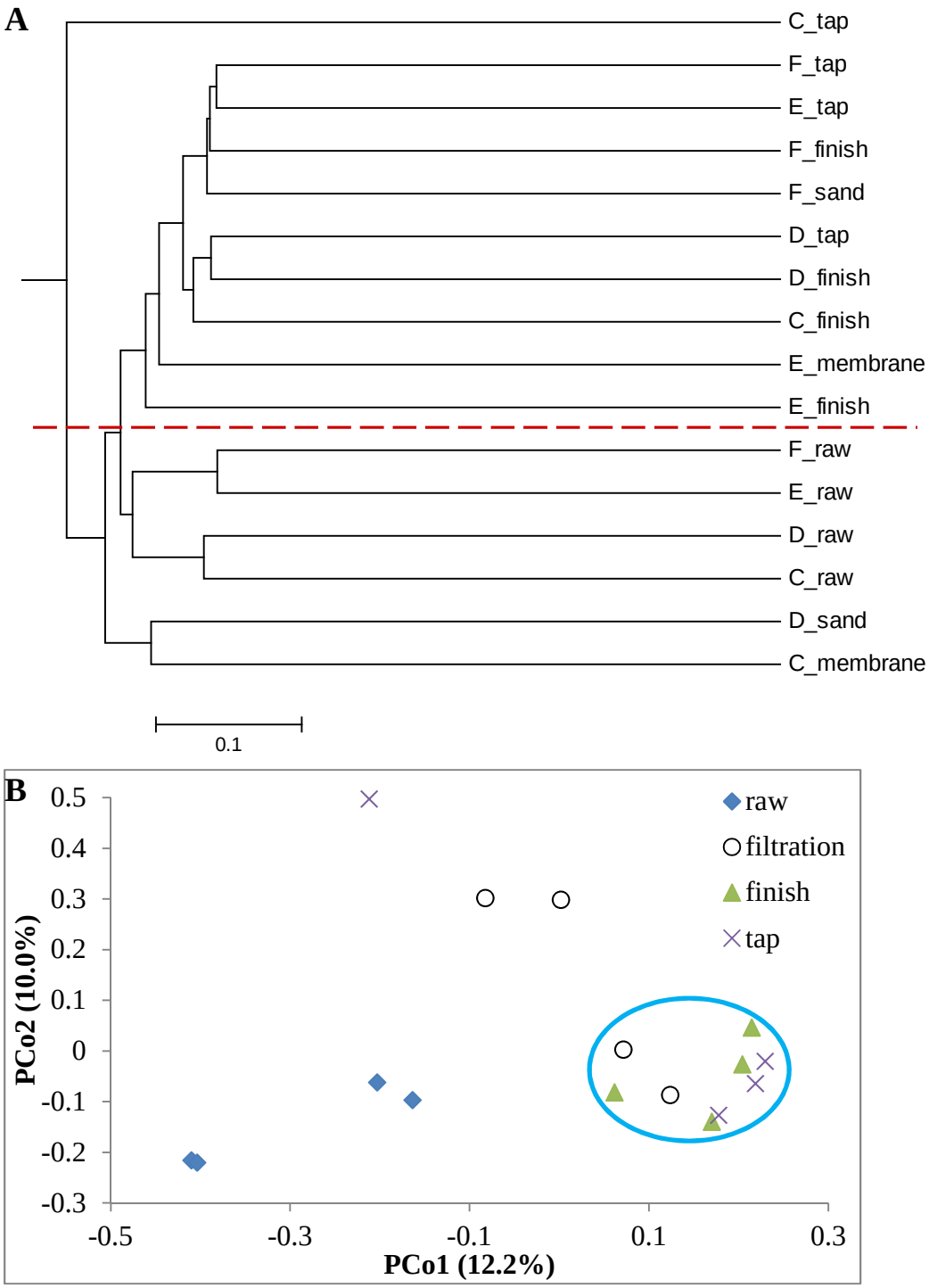

Figure 5.4 Hierarchical cluster analysis (A) and PCoA analysis (B) based on unweighted UniFrac distance matrix.

The corresponding PCoA analyses based on both weighed and unweighted analyses showed in Figure 5.3 B and 5.4 B. Consistent with the cluster analyses, in both PCoA analyses based weighed and unweighted UniFrac distance, the finished water, and tap water were grouped together suggesting that chlorination was a major factor that influenced the bacterial community structure. The group order was also similar to the hierarchical cluster tree: treatment processes > water sources > treatment technique. Distribution system also showed substantial effects on the tap water.

When the species relative abundance was considered, the two PCo axes explained total 38.3% variances (Figure 5.3 B). Both PCo axes reflected the separation caused by treatment processes as water samples from different treatment steps spread along the diagonal of the plot. Without considering the species relative abundance, PCo 1 reflected the separation caused by treatment processes, as water samples from different treatment steps mainly spread along PCo 1 axis which explained 12.2% total species variances. The tap water from water treatment plant C and the filtered water from plant C and D separated with other water samples along PCo 2 axis which accounted for 10.0% of total species variances.

Similar to our results Poitelon also found that chlorination is the major step that governed the bacterial community in drinking water distribution systems (Poitelon et al. 2010). Eichler's also conclude that the influence of unspecific coagulation, flocculation and sand filtration treatment on bacteria community structure was smaller than chlorination(Eichler et al. 2006). However, Pinto et al found that filtration step play an important role in shaping the bacterial community in drinking water (Pinto et al. 2012). Pinto's samples were taken in one treatment plant. Our samples, Piotelon and Eichler's samples were all taken from multiple treatment

streams therefore were more representative. Drinking water treatment researches were always case by case study. Many factors such as the giant variations in the raw water, different treatment strategy details could also cause the differences for the final conclusions in different studies.

5.5. Conclusions

Chlorination was the key step that controlling the bacterial community structure in drinking water. Membrane filtration had better treatment efficiency than sand filtration, however, didn't cause significantly different bacterial communities. The influence order on the water bacterial community were Disinfection> Water sources > Filtration technology. Persistent core bacteria population which survived each treatment and distribution steps and appeared in all the four water treatment plants was from the genus of *Sphingomonas*. Core bacterial populations appeared in all the finished water and tap water samples were belonged to the family of *Sphingomonadaceae, Comamonadaceae, Moraxellaceae,* and *Pseudomonadaceae,* which were classified to the bacteria genera of *Sphingomonas, Acinetobacter,* and *Pseudomonas*

5.6. References

Alspach, B., Adham, S., Cooke, T., Delphos, P., Garcia-Aleman, J., Jacangelo, J., Karimi, A., Pressman, J., Schaefer, J., Sethi, S. and Publicatic, A.S.P. (2008) Microfiltration and ultrafiltration membranes for drinking water. Journal American Water Works Association 100(12), 84-97.

APHA (2005) Standard methods for the examination of water and wastewater., American Public Health Association, Washington, D.C.

Bottino, A., Capannelli, C., Del Borghi, A., Colombino, M. and Conio, O. (2001) Water treatment for drinking purpose: ceramic microfiltration application. Desalination 141(1), 75-79.

Craun, G.F., Brunkard, J.M., Yoder, J.S., Roberts, V.A., Carpenter, J., Wade, T., Calderon, R.L., Roberts, J.M., Beach, M.J. and Roy, S.L. (2010) Causes of Outbreaks Associated with Drinking Water in the United States from 1971 to 2006. Clinical Microbiology Reviews 23(3), 507-528.

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.

Guo, H., Wyart, Y., Perot, J., Nauleau, F. and Moulin, P. (2010) Low-pressure membrane integrity tests for drinking water treatment: A review. Water Research 44(1), 41-57.

Henne, K., Kahlisch, L., Brettar, I. and Hofle, 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.

Hill, V.R., Kahler, A.M., Jothikumar, N., Johnson, T.B., Hahn, D. and Cromeans, T.L. (2007) Multistate evaluation of an ultrafiltration-based procedure for simultaneous recovery of enteric microbes in 100-liter tap water samples (vol 73, pg 4218, 2007). Applied and Environmental Microbiology 73(19), 6327-6327.

Ho, L., Braun, K., Fabris, R., Hoefel, D., Morran, J., Monis, P. and Drikas, M. (2012) Comparison of drinking water treatment process streams for optimal bacteriological water quality. Water Research 46(12), 3934-3942.

Huse, S.M., Welch, D.M., Morrison, H.G. and Sogin, M.L. (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environmental Microbiology 12(7), 1889-1898.

Jacangelo, J.G. and Watson, M. (2002) Control of microorganisms in drinking water, American Society of Civil Engineers, Reston, Virginia.

Lautenschlager, K., Boon, N., Wang, Y.Y., Egli, T. and Hammes, F. (2010) Overnight stagnation of drinking water in household taps induces microbial growth and changes in community composition. Water Research 44(17), 4868-4877.

Li, D., Li, Z., Yu, J., Cao, N., Liu, R. and Yang, M. (2010) Characterization of bacterial community structure in a drinking water distribution system during an occurrence of red water. Appl. Environ. Microbiol. 76(21), 7171-7180.

Mull, B. and Hill, V.R. (2009) Recovery and Detection of Escherichia coli O157:H7 in Surface Water, Using Ultrafiltration and Real-Time PCR. Applied and Environmental Microbiology 75(11), 3593-3597.

Nikkari, S., Lopez, F.A., Lepp, P.W., Cieslak, P.R., Ladd-Wilson, S., Passaro, D., Danila, R. and Relman, D.A. (2002) Broad-range bacterial detection and the analysis of unexplained death and critical illness. Emerging Infectious Diseases 8(2), 188-194.

Peter-Varbanets, M., Margot, J., Traber, J. and Pronk, W. (2011) Mechanisms of membrane fouling during ultra-low pressure ultrafiltration. Journal of Membrane Science 377(1-2), 42-53.

Pinto, A., Xi, C. and Raskin, L. (2012) Bacterial community structure in the drinking water microbiome is governed by filtration processes. Environmental Science & Technology.

Poitelon, J.B., Joyeux, M., Welte, B., Duguet, J.P., Prestel, E., Lespinet, O. and Dubow, M.S. (2009) Assessment of phylogenetic diversity of bacterial microflora in drinking water using serial analysis of ribosomal sequence tags. Water Research 43(17), 4197-4206.

Poitelon, J.B., Joyeux, M., Welte, 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 & Biotechnology 37(2), 117-128.

Polaczyk, A.L., Narayanan, J., Cromeans, T.L., Hahn, D., Roberts, J.M., Amburgey, J.E. and Hill, V.R. (2008) Ultrafiltration-based techniques for rapid and simultaneous concentration of multiple microbe classes from 100-L tap water samples. Journal of Microbiological Methods 73(2), 92-99.

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., Stres, B., Thallinger, G.G., Van Horn, D.J. and Weber, C.F. (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology 75(23), 7537-7541.

Walsh, M.E., Zhao, N., Gora, S.L. and Gagnon, G.A. (2009) Effect of coagulation and flocculation conditions on water quality in an immersed ultrafiltration process. Environmental Technology 30(9), 927-938.

Wang, H.Y., He, F.H., Zhi, Z. and Gu, P. (2011) The Effects of Pre-Treated Membrane Backwash Water on the Quality of Finished Water from a Membrane System. Separation Science and Technology 46(12), 1887-1897.

Wang, Q., Garrity, G.M., Tiedje, J.M. and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73(16), 5261-5267.

White, C., Tancos, M. and Lytle, D.A. (2011) Microbial Community Profile of a Lead Service Line Removed from a Drinking Water Distribution System. Applied and Environmental Microbiology 77(15), 5557-5561.

Chapter 6. Effects of Pipe Materials on Biofilm Bacterial Community Composition

6.1. Abstract

The effect of pipe materials on biofilm bacterial community composition were investigated through 16S rDNA based pyrosequencing. Biofilm on the surface of copper pipe was different from PVC and galvanized iron. Biofilm developed on the surface of copper pipes was dominated by *Alphaproteobacteria,* characterized by high abundance of *Methylobacteriaceae*, *Erythrobacteraceae*. Biofilm on the surface of PVC and galvanized iron were dominated by both *Alphaproteobacteria* and *Betaproteobacteria*. Half of bacterial populations in each pipe biofilm were found also persistent in other pipes. Nine core abundant bacteria families were found existed in biofilm of all pipe material surfaces were *Mycobacteriaceae*, *Caulobacteraceae*, *Bradyrhizobiaceae*, *Methylobacteriaceae*, *Erythrobacteraceae*, *Sphingomonadaceae*, *Burkholderiales*_incertae_sedis, *Comamonadaceae,* and *Oxalobacteraceae*.

6.2. Introduction

Biofilm acting as a reservoir for both pathogens and other oligotrophic bacteria was considered a potential risk for human health (Craun and Calderon 2001, Liu et al. 2012b). The presence of biofilm also depleted disinfection and accelerated disinfection decay (Kiene et al. 1998, LeChevallier et al. 1990). Many factors affect biofilm formation such as disinfection type and concentration, nutrient content, temperature and pipe material (LeChevallier et al. 1990, Momba et al. 2000). Previous study found that the bacteria growth was limited when dissolved organic carbon (DOC) < 1 mg/L, or assimilable organic carbon (AOC) < 50 ug/L. Temperature control during drinking water distribution was not practical. Monochloramine was found more efficient

than chlorine in biofilm control. An average 2.0 mg/L chloramine residual was necessary for sufficient biofilm bacteria removal. Biofilm in the drinking water distribution system may take years to be developed (Martiny et al. 2003). However once the biofilm was established, it is very difficult to be eliminated. Increasing disinfection concentration, switching disinfectant and systematically flushing system couldn't achieve satisfaction effect (LeChevallier et al. 1990).

Most previous studies on the effects of pipe materials only focused on pathogens and some culturable bacteria (Camper et al. 2003, Niquette et al. 2000). And others disinfection efficiency differences for biofilm developed on different material surface (LeChevallier et al. 1990, Williams et al. 2005). With the development and application of molecular techniques, a few studies for biofilm bacterial community composition were reported in recent years (Hong et al. 2010, Schmeisser et al. 2003). However most sample were from water meters, valves or inserted coupons which cannot represent the pipe materials in use. Biofilm from real distribution networks was available only when demolition or construction happened. Therefore the effects of pipe materials on biofilm bacterial composition were still not very clear and conclusive.

The aim of this study is to compare the biofilm bacterial community composition developed on different pipe materials. In Knox County local area, the water pipes used in premise plumbing following such a pattern: most commercial buildings use copper pipes; new houses use polyvinyl chloride (PVC) pipes; some old buildings still use galvanized iron for drinking water distribution. Therefore we set up pipe lines in laboratory using three different pipe materials galvanized iron, copper and PVC. Their effects on bacterial community composition were investigated using 16S rDNA based pyrosequencing.

6.3. Material and Methods

6.3.1. Pipeline setup and biofilm bacteria sampling

To simulate the biofilm developed in drinking water distribution networks, six plumbing pipes were built in parallel in the laboratory with two galvanized iron, copper, and polyvinyl chloride (PVC) pipelines served as duplicates. Each pipeline was connected to a submersible pump sitting in a covered bucket holding 15 L tap water as inlet reservoir. The outlet hose was as connected to the same bucket so that the water in the reservoir can be circulated by a submersible pump. Each pipe line and its connected reservoir were isolated and separate from other sets. All the systems were circulated 1 hour every day. Water flow rate through the pipes is adjusted to 1 L/min by ball valves. All the containers and water pipes were disinfected with 10% hypochlorous acid, washed with autoclaved deionized water before circulating with tap water. A long hydraulic retention time for the reservoir water was achieved by replacing 3.75 L water from each bucket with and the same volume of fresh chlorinated drinking water every week. The feed fresh water was from the same tap faucet in the University of Tennessee laboratory. The water quality background of the feed water was listed in Table 6.1. The inner diameter of the pipe is 0.5 inch, and the total length is 60 inch with six U-turns. At 120 day the valves were disassembled and biofilm developed inside pipe lines was scrapped using sterilized cotton swap and stored at -80 °C for DNA extraction. The six biofilm samples from copper, Galvanized iron, and PVC pipes were labeled as Cop1, Cop2, Giron1, Giron2, PVC1, and PVC2, respectively.

Table 6.1 Water quality parameters

pH	Turbidity (NTU)	DOC (mg/L)	Hardness (mg/liter as $CaCO_3$)	Conductivity (µS/cm)	Free Cl_2 (mg/liter)	Fluoride (mg/liter)	Chloride (mg/liter)	Sulfate (mg/liter)	Nitrate (mg/liter)
6.74-6.78	0.06-0.09	1.11-1.28	60-75	20-23	2.05-2.20	1.20-1.21	11.86-17.07	17.85-26.94	1.29-2.12

6.3.2. Pyrosequencing of 16S rDNA amplicons

The whole genome DNA was extracted using FastDNA spin kit for soil (MP Biomedicals, Santa Anna, CA). Pyrosequencing of the 16S rRNA gene amplicon libraries was performed to characterize the bacterial diversity in drinking water. For each DNA sample, amplicon libraries were generated with primers targeting the V3 hypervariable region of the 16S rRNA gene: 338F (5'-ACTCCTACGGGAGGCAGCAG-3') and 533R (5'-TTACCGCGGCTGCTGGCAC-3') (Huse et al. 2008). Barcode sequences unique to each sample were attached to both primers following a previously described sample tagging approach (Hamady et al. 2008). Polymerase Chain Reaction (PCR) amplification was performed with the FastStart High Fidelity PCR system in a total volume of 50 µL, containing 5 µL of FastStart High Fidelity Reaction Buffer with 1.8 mM $MgCl_2$, 200µM dNTPs, 0.4 µM forward and reverse primers, 10-100 ng of DNA, and 2.5 U FastStart High Fidelity Enzyme Blend (Roche Diagnostics, Germany). PCR was performed with the following thermal cycling program: 1 cycle at 94 °C for 3 min, 20 cycles at 94 °C for 30 sec, 57 °C for 1 min, 72 °C for 2 min, and a final extension at 72 °C for 2 min. Amplicons were purified with the Qiagen PCR purification kit (Qiagen, Valencia, California, USA) and the Agencourt AMPure PCR purification system (Beckman Coulter, Danvers, Massachusetts, USA). The quality of each amplicon library was evaluated using the Agilent DNA 7500 kit with a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, USA). Equal molar quantities of amplicons from each water sample were pooled together. The pooled DNA was immobilized

onto DNA capture beads and amplified through emulsion PCR using the GS FLX emPCR amplicon kit according to the manufacturer's protocols (454 Life Sciences, Branford, CT, USA). Sequencing of the PCR products was performed at the Center for Environmental Biotechnology at the University of Tennessee using a 454 Genome Sequencer FLX (454 Life Sciences, Branford, CT, USA).

6.3.3. Sequence analysis

Sequences acquired by pyrosequencing were parsed and trimmed according to sample-specific barcodes using GS Amplicon Variant Analyzer software version 2.3 (Roche Diagnostics, Germany). The analysis of 16S rRNA gene sequences obtained from pyrosequencing and clone library were both performed with MOTHUR program version 1.27.0 (Schloss et al. 2009). After denoising the flows and the removal of short sequences (<100 bp) the dataset was de-replicated to eliminate duplicate sequences. After checking and removing chimera sequences, the remaining sequences were de-replicated again and low-quality sequences were removed as described previously (Huse et al. 2010). The non-redundant sequences were aligned against the SILVA database alignment. A distance matrix was constructed for all subsequent microbial community ecological analyses, including community diversity and rarefaction analysis. Operational taxonomic units (OTUs) were assigned with the average neighbor clustering algorithm and taxonomy was assigned using the RDP classifier with a bootstrap cutoff of 80% (Wang et al. 2007). Multiple sample analysis was performed using UniFrac dendrograms.

Nucleotide sequence accession numbers

454 pyrosequencing raw data was submitted to NCBI GenBank Short Read Archive (SRA) with the accession number SRA036912.1.

6.4. Results and Discussion

6.4.1. Taxonomic composition of bacterial communities in biofilm

Bacteria community composition in biofilm was analyzed by 16S rDNA based high throughput pyrosequencing. Total of 45,793 reads with an average length of ~ 150 bp were generated from pyrosequencing. At 80% cut off level, 99.9% of those sequences were classified to known bacterial phyla. Total 7 bacterial phyla were identified: *Proteobacteria* (94.2%), *Bacteroidetes* (3.0%), *Actinobacteria* (2.5%), and *Planctomycetes* (0.1%), with less than 0.2% rare phyla composed of *Firmicutes, Acidobacteria,* and *Deinococcus-Thermus*. *Proteobacteria* was found the most predominate in all water samples with relative abundance above 91.0% in all biofilm samples. *Proteobacteria* was mainly affiliated to the class of *Alphaproteobacteria, Betaproteobacteria,* and *Gammaproteobacteria*.

At phylum level classification, biofilm on copper surface showed different bacterial community composition from biofilm developed on other pipe material surface (Figure 6.1 A). There was one most predominated population in copper biofilm *Alphaproteobacteria* accounted for 80.6% and 86.0% in two separate pipe systems. Biofilm on the surface of PVC and galvanize iron were dominated by both *Alphaproteobacteria* and *Betaproteobacteria*, both of which accounted for over 40% in PVC biofilm and one galvanized iron biofilm. In another galvanized iron biofilm the *Alphaproteobacteria* (71.4%) had high abundance than *Betaproteobacteria* (24.4%). *Gammaproteobacteria* in biofilm from all pipe lines were relative low with a range of 0.3% to 2.0%. One copper and one galvanized iron biofilm contained 6.9% *Actinobacteria*, which were further identified as *Mycobacterium* (6.3% and 4.6%). One PVC biofilm contained

7.6% *Bacteroidetes*, which were further classified into the genus of *Pedobacter* (2.8%) and *Flavobacterium* (1.1%). Henne studied biofilm on steel, copper, plastic, glass and Teflon surface from drinking water systems and found *Alphaproteobacteria* was the most dominant population (Henne et al. 2012). In their survey from single strand confirmation polymorphism (SSCP) fingerprints they detected 26% *Alphaproteobacteria*, 11% *Gammaproteobacteria*, and 9% *Betaproteobacteria*. Similar with our results Jang et al also found that *Alphaproteobacteria* dominant on both copper and PVC pipe surface based on DGGE analysis (Jang et al. 2011). Different from our observation, through T-RFLP analysis Pavissich concluded that *Betaproteobacteria* and *Gammaproteobacteria* were the dominant population on copper coupons (Pavissich et al. 2010). Schwartz found *Betaproteobacteria* (30 %) and *Gammaproteobacteria* (27%) were higher than *Alphaproteobacteria* (15%) on PVC surface (Schwartz et al. 2003). Kalmbach studied the effect of glass, low and high density polyethylene and soft PVC on bacterial community using FISH and found that *Betaproteobacteria* was the dominant population (Kalmbach et al. 2000). The different results may be caused by difference analysis methods.

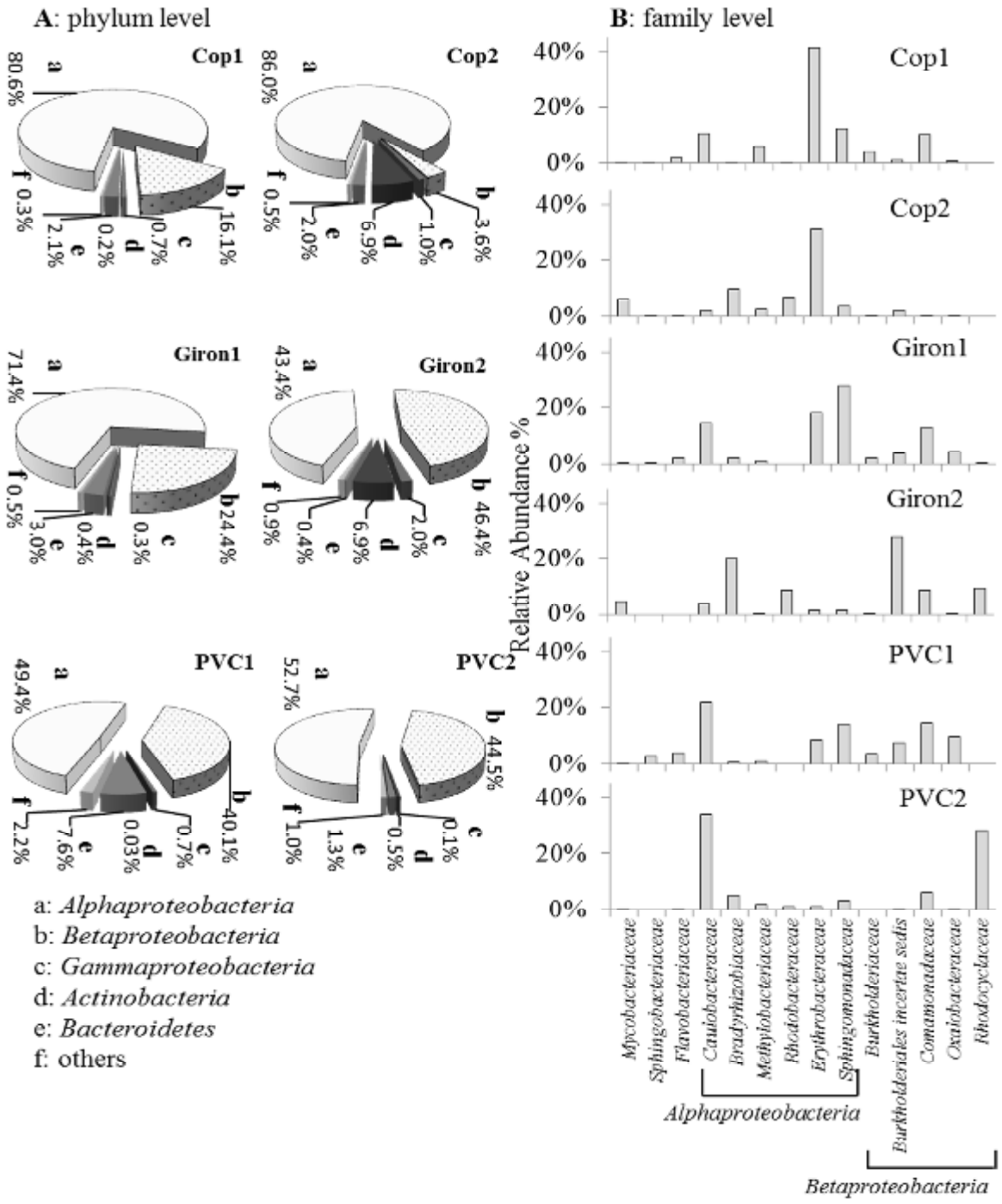

Figure 6.1 Biofilm bacterial community composition identified by 16S rRNA gene pyrosequencing from copper (Cop), galvanized iron (Giron) and polyvinyl chloride (PVC) pipe surface. Relative abundance at the phylum level, including classes of *Proteobacteria* with relative abundance > 1% are shown. Family level fingerprints were shown with relative abundance > 2%.

At family level classification, only 50 bacterial families were detected including 36 known bacteria families. The 36 families covered 89.0% and 67.6% copper, 91.8% and 90.8% galvanized iron and 91.2% and 83.9% PVC biofilm sequences. The abundant bacterial families (> 2% in at least one sample) which represented 64.7% ~ 90.4% total sequences from each biofilm samples were plotted in Figure 6.1 B. Differences among samples were obviously seen at family level of taxonomic classification. Biofilm in the two copper pipes were dominated by *Erythrobacteraceae* (41.5% and 31.2%). Most *Erythrobacteraceae* were not able to be classified at genus level. Only a small amount was identified as *Porphyrobacter* (1.3%). *Erythrobacteraceae* is a newly proposed family under the order of *Sphingomonadales* (Lee et al. 2005). The members of this family produce yellow, orange or pink pigment living in fresh or sea water habitats (Fuerst et al. 1993). *Caulobacteraceae, Methylobacteriaceae, Sphingomonadaceae* were also found in copper biofilm with the abundance ranged of 2.2% ~ 12.3%. The *Methylobacteriaceae* sequence was classified to the genus of *Methylobacterium* (5.7% and 2.5%). *Methylobacterium* was also detected and isolated from corroded copper pipe (Keevil 2004, Pavissich et al. 2010). *Methylobacterium* was reported ubiquitous in biofilm on other pipe material surface (Jang et al. 2011, White et al. 2011). In this study we found *Methylobacterium* persistent in all pipe with the highest relative abundance appeared in copper pipes (2.5% and 5.7%).

The most dominant population in the two PVC pipe biofilm was *Caulobacteraceae* (21.9% and 34.1%). Most *Caulobacteraceae* were not able to be classified at genus level. Only a small amount was identified as *Brevundimonas* (2.1% and 0.02%). *Brevundimonas* as a well-known aquatic origin bacterium was observed and isolated from biofilm of PVC and other plastic pipes (Silbaq 2009, Yu et al. 2010). The two galvanized iron pipe had different bacterial community

composition. One was dominated by *Sphingomonadaceae* (27.8%), *Erythrobacteraceae* (18.3%), *Caulobacteraceae* (14.6%), and *Comamonadaceae* (13.1%). Another mainly was composed of *Burkholderiales*_incertae_sedis (27.7%), *Bradyrhizobiaceae* (20.0%), and *Rhodocyclaceae* (9.6%), which were classified into the genus of *Aquabacterium* (27.5%), *Bradyrhizobium* (16.3%), and *Azospira* (9.5%), respectively. *Aquabacterium* existed inside all pipes with variable relative abundance: 1.1% and 2.2% in copper pipes, 3.8% and 27.5% in Galvanized iron and 7.6% and 0.05% in PVC pipes. Kalmbach observed *Aquabacterium* dominated in biofilm on glass, low and high density polyethylene and soft PVC pipe surface (Kalmbach et al. 2000). *Bradyrhizobium* was found the dominant genus in ductile cast iron pipe by other researchers (Jang et al. 2012). *Azospira* was also aquatic origin bacteria and recently been observed in tap faucet gasket (Bae et al. 2007, Liu et al. 2012a).

6.4.2. Biofilm bacterial comparison developed inside different pipe lines

To compare the bacterial communities composition in biofilm developed on different pipe materials, hierarchical cluster analyses were performed at family level classification based on weighted UniFrac distance (Figure 6.2). On both trees the copper biofilm were clustered together, suggesting that biofilm developed on copper surface were different from biofilm on other material surface. Since the abundant families (> 2%) represented most sequences for each biofilm sample (87.8% and 64.7% for copper; 90.4% and 86.0% for galvanized iron; 89.2% and 81.6% for PVC), principal component analysis was performed surface based on abundance bacterial families to investigate the variances of bacterial community composition in different biofilm (Figure 6.3). The two principle component represented total 71% variances. Biofilm developed on the surface of copper pipes were separate with PVC and galvanized iron due to the bacterial families of *Methylobacteriaceae*, *Erythrobacteraceae.*

6.4.3. Core population and pathogen signature in biofilm

Taxonomic analysis revealed a broad range of bacteria in biofilm on different pipes surface. Total 50 bacteria families and 91 genera were detected in the biofilm. A lot of core populations were found persistent in all water samples. Total 20 families and 22 genera were shared by all biofilm samples. Nine core abundant bacteria families (2%) were found persist in biofilm of all pipe material surfaces: *Mycobacteriaceae*, *Caulobacteraceae*, *Bradyrhizobiaceae*, *Methylobacteriaceae*, *Erythrobacteraceae*, *Sphingomonadaceae*, *Burkholderiales*_incertae_sedis, *Comamonadaceae*, and *Oxalobacteraceae*. These core abundant bacteria accounted for over half of total population in every biofilm samples (Figure 6.4). From copper to PVC, the relative abundance of *Erythrobacteraceae* was decreased, *Methylobacteriaceae* was increase, and *Comamonadaceae* was more constant in each sample. Except for *Erythrobacteraceae*, most of these bacterial families were found ubiquitous in finished water and tap water from our previous study.

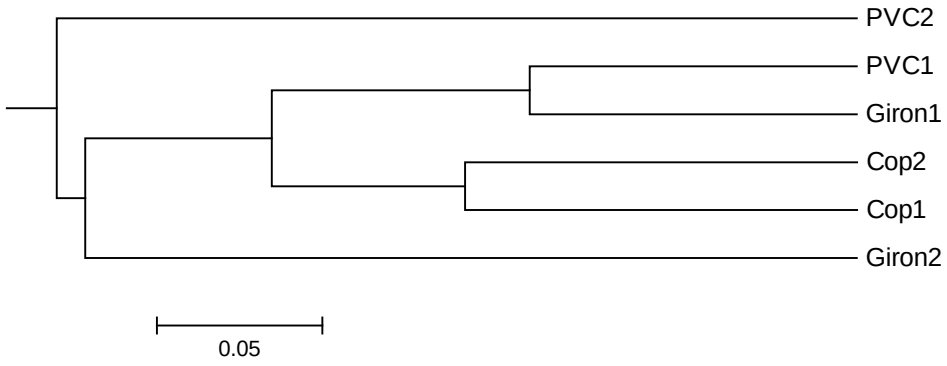

Figure 6.2 Hierarchical cluster analysis based on weighted UniFrac distance matrix. Cop: copper; Giron: galvanized iron; PVC: polyvinyl chloride

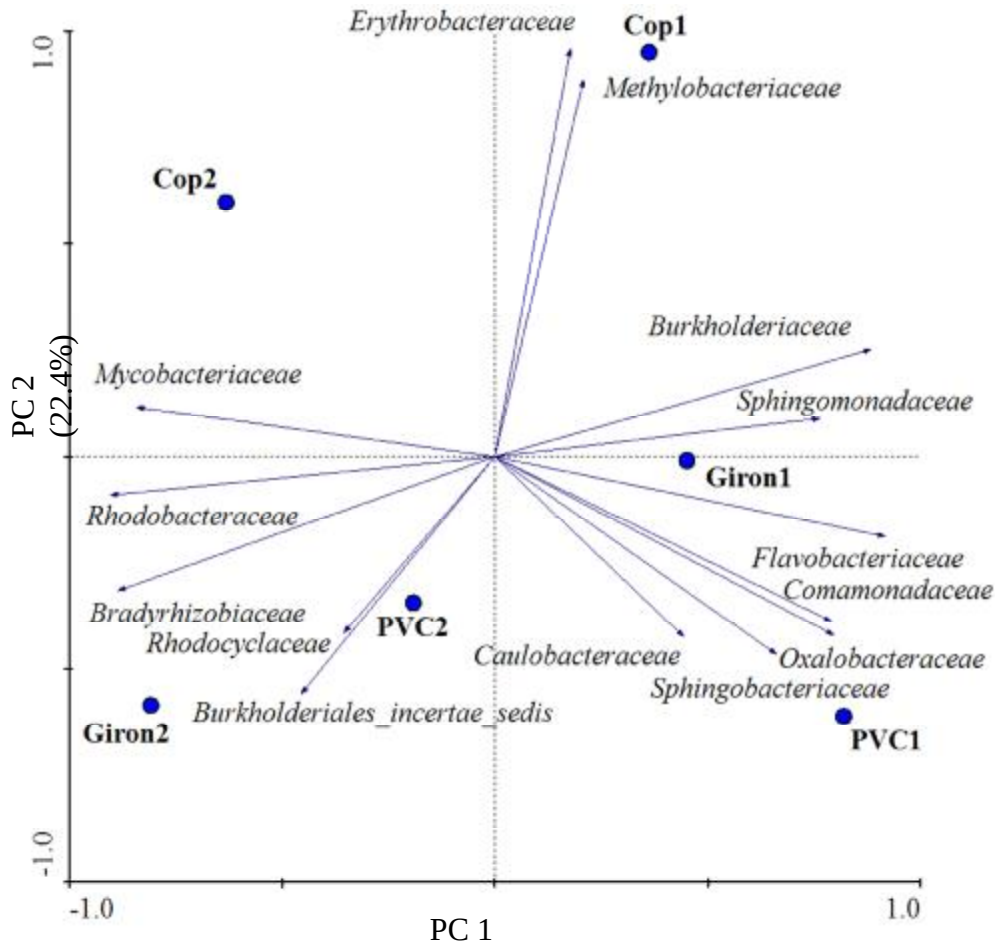

Figure 6.3 Principal component analysis (PCA) of biofilm bacterial communities at family level with relative abundance above 2%. Every vectors point to the direction of increase for a given variable so that biofilm samples with similar communities are located in the similar positions in the diagram. Cop: copper; Giron: galvanized iron; PVC: polyvinyl chloride

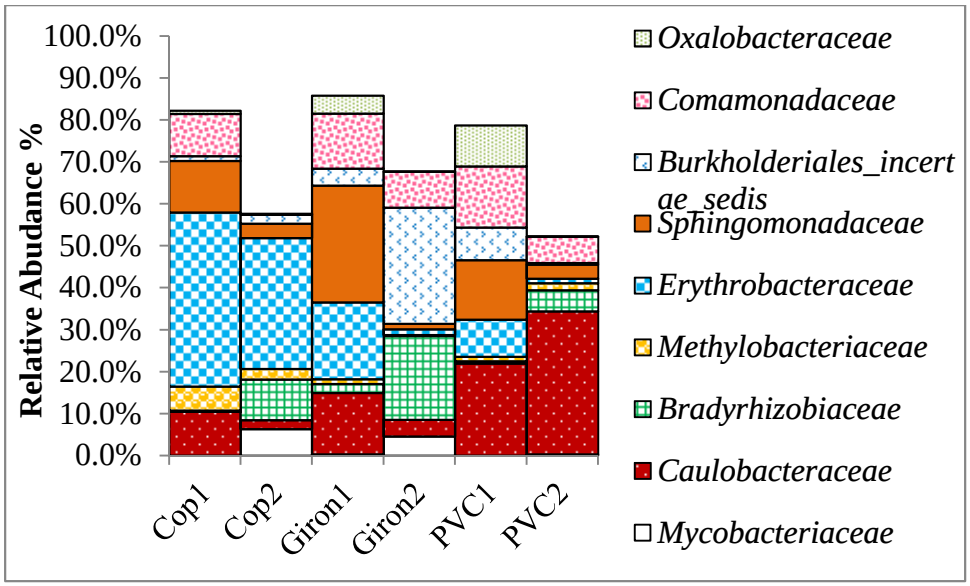

Figure 6.4 Core abundant bacteria families (> 2%) existed in biofilm of all pipe material surfaces. Cop: copper; Giron: galvanized iron; PVC: polyvinyl chloride

Two pathogens were detected in the biofilm samples: *Mycobacterium* and *Legionella*. Mycobacterium was found in all pipes (0.1% and 6.3% in copper, 0.3% and 4.6% in galvanized iron, 0.01% and 0.3% in PVC). *Legionella* only detected in one copper and one galvanized iron pipe with low content (0.9% and 1.6%). *Mycobacterium* and *Legionella* were observed in biofilm of drinking water distribution systems with high frequency by other researchers (Liu et al. 2012b, Schwartz et al. 1998, Schwartz et al. 2003).

Copper and other metallic pipes were reported caused different bacterial community compared than plastic pipes due to the metal ion on the surface. Many researchers observed the community differences between biofilm on metallic and plastic surface (Jang et al. 2011, Schwartz et al. 2003). Other study didn't found significant differences caused by pipe materials (Henne et al. 2012, Zacheus et al. 2000). Biofilm formation was considered a slow process. Martiny's study showed that biofilm may take years to be developed in the drinking water distribution system (Martiny et al. 2003). Henne's results indicated that young biofilm was affect

more by the pipe material, during years of maturity, the biofilm communities will show similar structure as their physically related neighbors (Henne et al. 2012).

6.5. Conclusions

Biofilm developed on the surface of copper pipes was dominated by *Alphaproteobacteria*. Biofilm on the surface of PVC and galvanize iron were dominated by both *Alphaproteobacteria and Betaproteobacteria*. Biofilm developed on the surface of copper pipe was different than PVC and galvanized iron, characterized by high abundance of *Methylobacteriaceae, Erythrobacteraceae.* Half of bacterial populations in each pipe biofilm were found also persistent in other pipes. These core abundant bacteria families existed in biofilm of all pipe material surfaces were *Mycobacteriaceae, Caulobacteraceae, Bradyrhizobiaceae, Methylobacteriaceae, Erythrobacteraceae, Sphingomonadaceae, Burkholderiales_incertae_sedis, Comamonadaceae,* and *Oxalobacteraceae.*

6.6. References

Bae, H.-S., Rash, B.A., Rainey, F.A., Nobre, M.F., Tiago, I., da Costa, M.S. and Moe, W.M. (2007) Description of Azospira restricta sp. nov., a nitrogen-fixing bacterium isolated from groundwater. International Journal of Systematic and Evolutionary Microbiology 57(7), 1521-1526.

Camper, A.K., Brastrup, K., Sandvig, A., Clement, J., Spencer, C. and Capuzzi, A.J. (2003) Effect OF DISTRIBUTION SYSTEM MATERIALS on bacterial regrowth. Journal (American Water Works Association) 95(7), 107-121.

Craun, G.E. and Calderon, R.L. (2001) Waterborne disease outbreaks caused by distribution system deficiencies. Journal American Water Works Association 93(9), 64-75.

Fuerst, J.A., Hawkins, J.A., Holmes, A., Sly, L.I., Moore, C.J. and Stackebrandt, E. (1993) Porphyrobacter neustonensis gen. nov., sp. nov., an Aerobic Bacteriochlorophyll-Synthesizing Budding Bacterium from Fresh Water. International Journal of Systematic Bacteriology 43(1), 125-134.

Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J. and Knight, R. (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nature Methods 5(3), 235-237.

Henne, K., Kahlisch, L., Brettar, I. and Hofle, 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.

Hong, P.Y., Hwang, C.C., Ling, F.Q., Andersen, G.L., LeChevallier, M.W. and Liu, W.T. (2010) Pyrosequencing Analysis of Bacterial Biofilm Communities in Water Meters of a Drinking Water Distribution System. Applied and Environmental Microbiology 76(16), 5631-5635.

Huse, S.M., Dethlefsen, L., Huber, J.A., Welch, D.M., Relman, D.A. and Sogin, M.L. (2008) Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. Plos Genetics 4(11).

Huse, S.M., Welch, D.M., Morrison, H.G. and Sogin, M.L. (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environmental Microbiology 12(7), 1889-1898.

Jang, H.J., Choi, Y.J. and Ka, J.O. (2011) Effects of Diverse Water Pipe Materials on Bacterial Communities and Water Quality in the Annular Reactor. Journal of Microbiology and Biotechnology 21(2), 115-123.

Jang, H.J., Choi, Y.J., Ro, H.M. and Ka, J.O. (2012) Effects of Phosphate Addition on Biofilm Bacterial Communities and Water Quality in Annular Reactors Equipped with Stainless Steel and Ductile Cast Iron Pipes. Journal of Microbiology 50(1), 17-28.

Kalmbach, S., Manz, W., Bendinger, B. and Szewzyk, U. (2000) In situ probing reveals Aquabacterium commune as a widespread and highly abundant bacterial species in drinking water biofilms. Water Research 34(2), 575-581.

Keevil, C.W. (2004) The physico-chemistry of biofilm-mediated pitting corrosion of copper pipe supplying potable water. Water Science and Technology 49(2), 91-98.

Kiene, L., Lu, W. and Levi, Y. (1998) Relative importance of the phenomena responsible for chlorine decay in drinking water distribution systems. Water Science and Technology 38(6), 219-227.

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

Lee, K.-B., Liu, C.-T., Anzai, Y., Kim, H., Aono, T. and Oyaizu, H. (2005) The hierarchical system of the 'Alphaproteobacteria': description of Hyphomonadaceae fam. nov., Xanthobacteraceae fam. nov. and Erythrobacteraceae fam. nov. International Journal of Systematic and Evolutionary Microbiology 55(5), 1907-1919.

Liu, R., Yu, Z., Guo, H., Liu, M., Zhang, H. and Yang, M. (2012a) Pyrosequencing analysis of eukaryotic and bacterial communities in faucet biofilms. Science of The Total Environment 435–436(0), 124-131.

Liu, R., Yu, Z., Zhang, H., Yang, M., Shi, B. and Liu, X. (2012b) Diversity of bacteria and mycobacteria in biofilms of two urban drinking water distribution systems. Canadian Journal of Microbiology 58(3), 261-270.

Martiny, A.C., Jorgensen, T.M., Albrechtsen, H.J., Arvin, E. and Molin, S. (2003) Long-term succession of structure and diversity of a biofilm formed in a model drinking water distribution system. Applied and Environmental Microbiology 69(11), 6899-6907.

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

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.

Pavissich, J.P., Vargas, I.T., González, B., Pastén, P.A. and Pizarro, G.E. (2010) Culture dependent and independent analyses of bacterial communities involved in copper plumbing corrosion. Journal of Applied Microbiology 109(3), 771-782.

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., Stres, B., Thallinger, G.G., Van Horn, D.J. and Weber, C.F. (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology 75(23), 7537-7541.

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.

Schwartz, T., Kalmbach, S., Hoffmann, S., Szewzyk, U. and Obst, U. (1998) PCR-based detection of mycobacteria in biofilms from a drinking water distribution system. Journal of Microbiological Methods 34(2), 113-123.

Schwartz, T., Hoffmann, S. and Obst, U. (2003) Formation of natural biofilms during chlorine dioxide and u.v. disinfection in a public drinking water distribution system. Journal of Applied Microbiology 95(3), 591-601.

Silbaq, F.S. (2009) Viable ultramicrocells in drinking water. Journal of Applied Microbiology 106(1), 106-117.

Wang, Q., Garrity, G.M., Tiedje, J.M. and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73(16), 5261-5267.

White, C., Tancos, M. and Lytle, D.A. (2011) Microbial Community Profile of a Lead Service Line Removed from a Drinking Water Distribution System. Applied and Environmental Microbiology 77(15), 5557-5561.

Williams, M.M., Domingo, J.W.S. and Meckes, M.C. (2005) Population diversity in model potable water biofilms receiving chlorine or chloramine residual. Biofouling 21(5-6), 279-288.

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.

Zacheus, O.M., Iivanainen, E.K., Nissinen, T.K., Lehtola, M.J. and Martikainen, P.J. (2000) Bacterial biofilm formation on polyvinyl chloride, polyethylene and stainless steel exposed to ozonated water. Water Research 34(1), 63-70.

Chapter 7. Application of Zipf-Mandelbrot Model to Drinking Water Bacterial Community Distribution

7.1. Abstract

Relative abundance distribution (RAD) is an important way to illustrate the community structure. Bacterial community structure in drinking water was poorly characterized. Bacterial community composition in drinking water and biofilm developed in lab scale pipe lines were investigated using pyrosequencing analysis. Zipf-Mandelbrot (ZM) model was used to quantify the relative abundance distribution of bacterial communities. About 90% of the total variances for both bulk water and biofilm bacteria community structure were explained by ZM model, which indicated that the bacterial communities were all characterized by a few very abundant species followed by a long tail of rare species. Drinking water and the surface of pipes showed low niche diversity as reflected by the low β value of the bacteria RADs. Biofilm developed on the same pipe materials had similar RAD shapes as reflected by the similarity of both γ and β values, suggesting that pipe material may play an important role in the biofilm bacterial community assembly.

7.2. Introduction

Bacterial communities display various structural patterns in both natural and engineered environments (Inceoglu et al. 2011, Schloss and Handelsman 2006, Sloan et al. 2007). Relative abundance distribution (RAD) and species abundance distribution (SAD) are the two major ways to illustrate community structure (McGill et al. 2007). The RAD is widely used to compare the structure differences since the SAD curves are considered biased by the arbitrary abundance categories (Wilson 1991). When a large number of more abundant species and a fewer number of

rare species coexisted in the community the RAD will show as a convex curve. With a fewer number of more abundant species and a large number of rare species a community will show a concave curve. A community with more abundant and rare but less species in the middle will show an invert S shaped curve. Many models were developed to quantify the RAD in the past 70 years (McGill et al. 2007). For example, the linear shape RAD was usually described use geometric series (Narang and Dunbar 2004). The lognormal and broken stick model had better fit for the invert S shaped curves (Dunbar et al. 2002, Wilson 1991). Power law and Zipf model resulted in concave curves (Frontier 1985, Inceoglu et al. 2011). Except for the observation and description of RAD, all researches expected more ecological information and sought an answer to the question about how this specific community structure was shaped. Ecologists generalized all kinds of RAD patterns and developed more than 40 ecological models to integrate RAD with ecological processes. The rationale behind the ecological models is that RAD patterns are formed from specific community assembly processes; and therefore the RADs might indicate the corresponding process. The mathematical model derived from the ecology process might predict a particular community structure. One possible way to link the observed RAD to its ecological context is to find the best fit ecological model for the distribution curve and explain the community assembly process through the best fit model.

Most ecological models were extensively tested and used in macro ecology to address the ecological rules that may govern the assembly of plant and animal communities (Hubbell 2001, Watkins and Wilson 1994). With the development of molecular techniques the whole community, instead of only the culturable microbes, was able to be sampled and investigated. Till very recent years, microbial ecologists managed to extrapolate some models from macro ecological to micro ecology (Horner-Devine et al. 2004, Prosser et al. 2007). The development of next generation

sequencing significantly improved the coverage of bacterial community in environmental samples, particularly rare populations that could not be readily identified by other techniques. The massive information allowed us to test the macro ecological theory in micro scale. A few studies were focus on the bacterial community structures in soil ocean and even wastewater (Galand et al. 2009, Schloss and Handelsman 2006, Sloan et al. 2007). However, the model tests on drinking water community were very rare.

Bacteria and biofilm in drinking water are considered potential risks for human health (Craun et al. 2010, Henne et al. 2012). Most studies only focused on the total bacterial numbers, specific pathogens and the effects of disinfectant (Hammes et al. 2008, Poitelon et al. 2010, Wang et al. 2012). However the whole picture of bacterial community is not well characterized. In our recent studies, bacterial community structures in drinking water all appeared as concave curves. Therefore we selected a typical concave shaped model Zipf-Mandelbrot model to characterize the bacteria relative abundance distributions. The Zipf-Mandelbrot model (ZM) was first introduced to assess the information cost (Frontier 1985). In recent years the ZM model has been gradually applied in ecology, genomics, and metabolic pathway distributions (Almaas et al. 2004, Kuznetsov 2003, Wilson et al. 1996). In the current study we investigate bacterial community in drinking water and biofilm on lab-scale pipeline interior surface though 16S rDNA based pyrosequencing. Zipf-Mandelbrot model was tested for fitness and used to quantitatively characterize bulk water and biofilm bacterial community structure at species level. However the conjectures on the ecological meanings required further experimental validation.

7.3. Material and Methods

7.3.1. Bacteria sampling from bulk drinking water and biofilm

In order to investigate the bacterial community assembly pattern in drinking water, bacteria in both bulk water and biofilm were sampled. Bulk water was collected from five taps affiliated to five different water treatment plants in Knox County Tennessee USA in June 2010. The bulk waters were marked as 1F, 2R, 3S, 4L and 5P, respectively. The first four tap waters were supplied by water treatment plants use conventional treatment process including coagulation, flocculation, sedimentation, sand filtration, and disinfection with chlorine. The tap water 5P is from a membrane filtration plant. The raw water is pumped through submerged microfiltration Siemens membranes after coagulation and flocculation, and then a disinfection with chlorine served as the last step. Before sampling, tap water faucet was flushed for 10 min at maximum flow rate. Total 150 liters of water from each tap were collected with autoclaved polyethylene carboys and transported into laboratory. Bacteria were harvested within four hours after water sampling. The drinking water bacteria were concentrated with a tangential-flow ultrafiltration system configured, prepared, and operated as previously described (Polaczyk et al. 2008). The biofilm samples were collected from pipe lines setup in the laboratory to simulate drinking water distribution networks. Biofilm developed on three different pipe materials Galvanized iron, Copper and PVC were sampled using sterilized cotton swap and stored at -80 °C for DNA extraction. Total six biofilm samples were collected from duplicate pipe lines for each pipe material. The biofilm samples from copper, Galvanized iron, and PVC pipes were marked as Cop1, Cop2, Giron1, Giron2, PVC1, and PVC2, respectively.

7.3.2. Bacterial community analysis and relative abundance distribution (RAD)

The whole genome DNA from both bulk water and biofilm were extracted using FastDNA spin kit for soil (MP Biomedicals, Santa Anna, CA). The bacterial communities were analyzed by Pyrosequencing of the 16S rRNA gene amplicon libraries generated from each samples as described before (Huse et al. 2008); (Nikkari et al. 2002). Sequencing of the amplicons was performed at the Center for Environmental Biotechnology at the University of Tennessee using 454 Genome Sequencer FLX titanium platform (454 Life Sciences, Branford, CT, USA). Sequences generated by pyrosequencing were processed using MOTHUR program version 1.27.0 (Schloss et al. 2009). The relative abundance distribution was calculated after denoising, screening, removing the chimeras and bad quality sequences follow MOTHUR pipeline. The high quality sequences were aligned and clustered with the average neighbor method to form Operational taxonomic units (OTUs).

7.3.3. Relative abundance distribution (RAD) and Zipf-Mandelbrot (ZM) model

Community structure is usually illustrated by the relative abundance of species ranked from high abundance to low abundance. The RAD pattern usually reflects certain ecological processes which play a major role in shaping the community structure. To simulate the species distribution pattern, pyrosequencing OTUs with a cut off of 3% dissimilarity were used for the RAD plot. Since all the RADs were concave curves, a typical concave shaped model, Zipf-Mandelbrot model were used to describe relative abundance distribution of bacteria species as previously described (Frontier 1985, Magurran 2004). As shown in Eq.1 parameters: F_0, β and γ were changed iteratively to minimize the sum of square of observed data the simulated data.

$$F_i = F_0(i+\beta)^{-\gamma} \quad \text{Eq.1}$$

Where F_i is the relative abundance of *i*th species; F_0 ($0 < F_0 < 1$) is the fitted relative abundance of the most abundant species so that sum of F_i equates to 1; S is the number of observed species; i ($i = 1, 2, \ldots, S$) is ranked order; β ($\beta > -1$) is the potential diversity of the environment; γ ($\gamma > 1$) is the average possibility of the appearance of a species.

7.3.4. Model fitting of Relative abundance distribution (RAD)

Since relative abundance curve was usually plot in logarithm scale, the Zipf-Mandelbrot model was also expressed in logarithmic form:

$$\log F_i = \log F_0 - g \log(i + \beta) \quad \text{Eq.2}$$

Parameters F_0, β and γ were changed iteratively to minimize the sum of squared residuals. For the best fit, R^2 ($1\text{-}SS_{model}/SS_{total}$) was calculated to evaluate how much variances can be explained by the model with optimized parameters.

To investigate the influence of model parameter changes on the community diversity, the Shannon diversity index (*H'*) and evenness index (*E*) were calculated for each sample (Eq.3 and Eq.4).

$$H' = -\sum_{i=1}^{S} (F_i \ln F_i) \quad \text{Eq.3}$$

$$E = \frac{H'}{\ln S} \quad \text{Eq.4}$$

7.4. Results and Discussion

7.4.1. Fit Zipf-Mandelbrot model to bulk tap water and biofilm bacterial communities

Relative abundance distributions (RADs) were plotted for all the water samples using ranked OTUs with 97% similarities to simulate the species abundance. The total number of OTUs

identified in bulk water samples and biofilm fell in a similar range (between 189 and 345 for bulk water, between 210 and 349 for biofilm samples). Fitting of the observed relative distribution showed that Zipf-Mandelbrot model fit all the RAD patterns and explained about 90% variances of RADs for all the samples (Figure 7.1 and 7.2). The best fit parameters and the R^2 values for goodness of fit were summarized in Table 7.1. There were no significant differences for the γ values in bulk water and biofilm samples with a range from 1.23 to 1.48. The parameter γ represents the average possibility of the occurrence of species. The value near 1 gives a higher evenness for the community (Frontier 1985, Wilson et al. 1996). The β values for both bulk water and biofilm were less than 0 with a range from -0.84 to -0.34. The parameter β represents the degree of niche diversification. The low β value indicated low niche diversity in both drinking water and biofilm.

Table 7.1 Summary of the best fit Zipf-Mandelbrot parameters and the diversity and evenness of the bacterial community

	γ	β	R^2	Shannon	Evenness
1F	1.39	-0.34	0.95	2.99	0.54
2R	1.27	-0.60	0.93	2.81	0.50
3S	1.32	-0.64	0.95	2.53	0.45
4L	1.24	-0.54	0.94	3.27	0.56
5P	1.45	-0.74	0.90	2.09	0.40
Cop1	1.25	-0.84	0.93	2.13	0.38
Cop2	1.23	-0.83	0.89	2.52	0.43
Giron1	1.35	-0.60	0.94	2.74	0.49
Giron2	1.36	-0.52	0.94	2.83	0.52
PVC1	1.41	-0.44	0.92	3.13	0.53
PVC2	1.48	-0.44	0.95	2.62	0.49

Zipf-Mandelbrot model assumes that a species is dependent on previous conditions and previous existing species (Frontier 1985, McGill et al. 2007). Pioneer species requiring low cost and prior conditions to invade, and subsequently will became the abundant species in the community. The later invaders and late succession species need high cost for resources and energy to invade the environment. They can invade the community only when the necessary conditions were met. And therefore the late succession species are usually rare species. Therefore, when many environmental factors work sequentially on species, the relative abundance pattern will appear as a Zipf-Mandelbrot distribution. Some botanists use this model to explain the assembling of plant communities (Watkins and Wilson 1994). For bacteria community patterns some researchers reported the best fit of power law for bacterial communities in soil (Inceoglu et al. 2011). Others use exponential regression (Henne et al. 2012). Basically both power law and Zipf-Mandelbrot model yield similar model plots. Power law model is similar as Zipf model, a special case of Zipf-Mandelbrot when β is 0. As the subset of the Zipf-Mandelbrot model, neither of them will fit better than Zipf-Mandelbrot model.

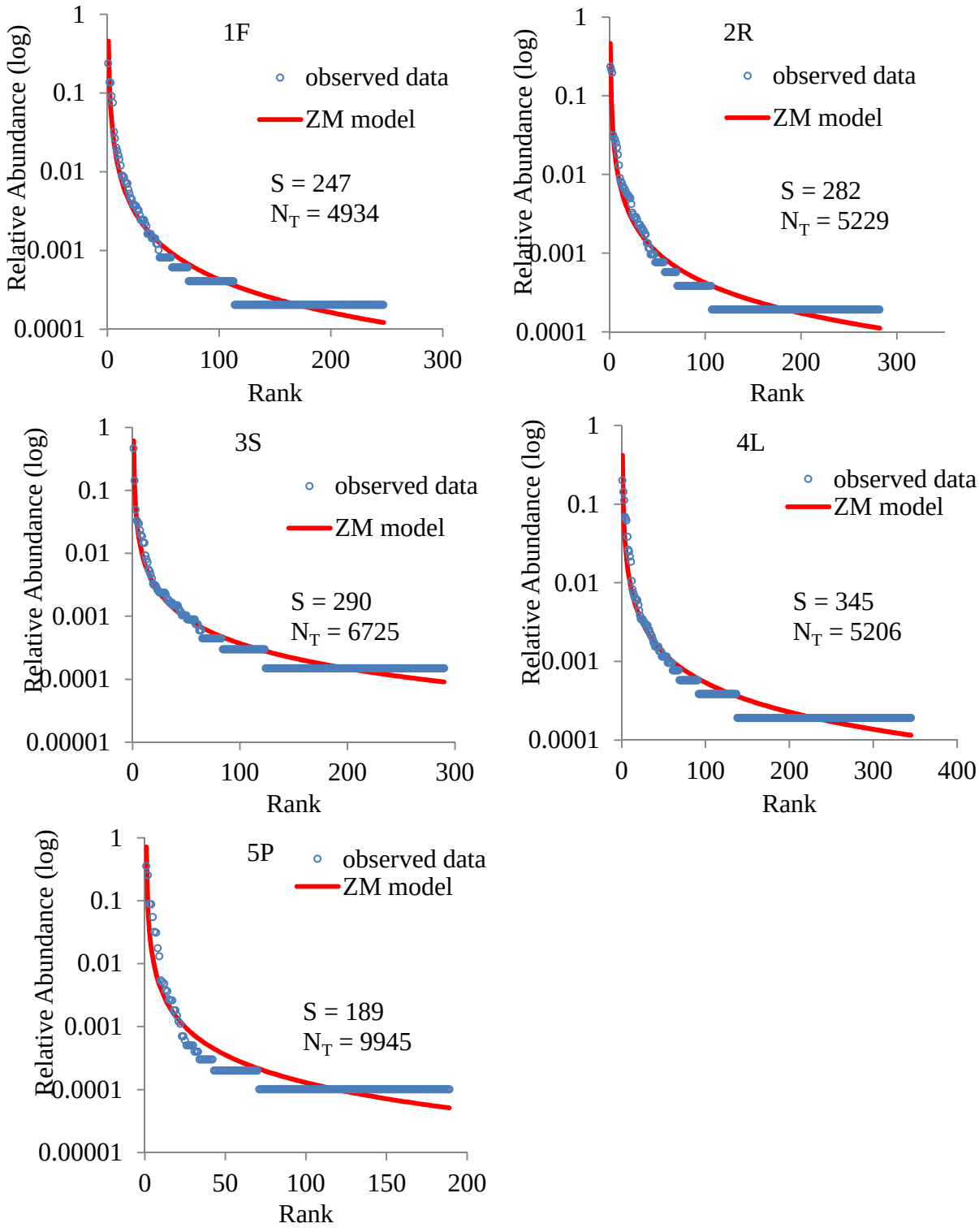

Figure 7.1 Relative abundance distribution (RAD) for bacterial community in drinking water. Zipf-Mandelbrot (ZM) model was fitted to the data. S is the total number of OTUs. N_T is the total number of sequences.

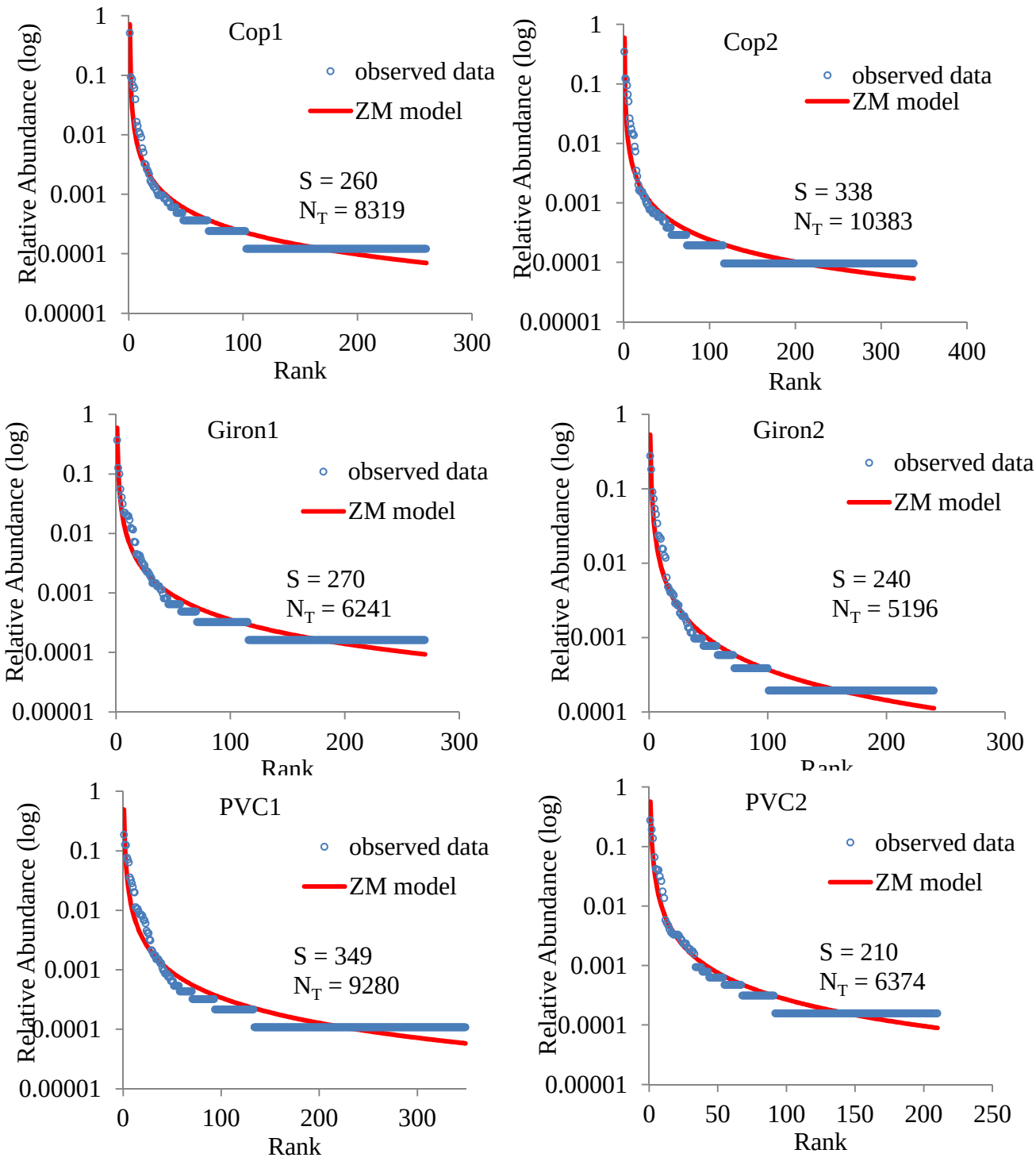

Figure 7.2 Relative abundance distribution (RAD) for bacterial community on different pipe material surface. Zipf-Mandelbrot (ZM) model was fitted to the data. S is the total number of OTUs. N_T is the total number of sequences.

7.4.2. Effects of model parameters on the shape the relative abundance curve and community diversity

Relative abundance distributions based on Zipf-Mandelbrot model were generated to test the effects of parameters on the curve shape. Frontier suggested that γ values are usually not greater than 2 ~ 4 and rarely less than 1 in ecology (Frontier 1985). Wilson expanded the γ value up to 10 to describe the plant communities (Wilson 1991). The β values usually ranged between -1 and 5. To simulate the ecological communities that were close to our drinking water sample, the upper bounds for the two parameters were both set for 5 in this study. Six γ and β pairs were selected to represent the relative abundance distributions in appropriate ecological ranges (Figure 7.3). For reasonable comparison, the simulations were performed based on the same total species richness (total OTU number in sample 1F). As shown in Figure 7.3 A and 7.3 C, a large γ value results a deep concave curve, while a large β value gives a shallow concave curve. When the rank of species was also plotted in log scale, the relative abundance distributions are close to straight lines (Figure 7.3 B and 7.3 D). The γ value controls the slop which represents the average abundance of all species. An increase of γ value results a deep slope. The β value controls the head of the slope which represents the evenness of the abundant species. When β is equal to 0, the log-log plot of RAD is a straight line. When β is larger than 0, the head of the slop will be shallower. The abundant species will be distributed more evenly. When β is smaller than 0, the head of the slop is deeper. The abundant species will be distributed less evenly. In our case a steep slope and lower head evenness was observed for all bacterial community RADs in both bulk water and biofilm samples as the β values were all smaller than 0.

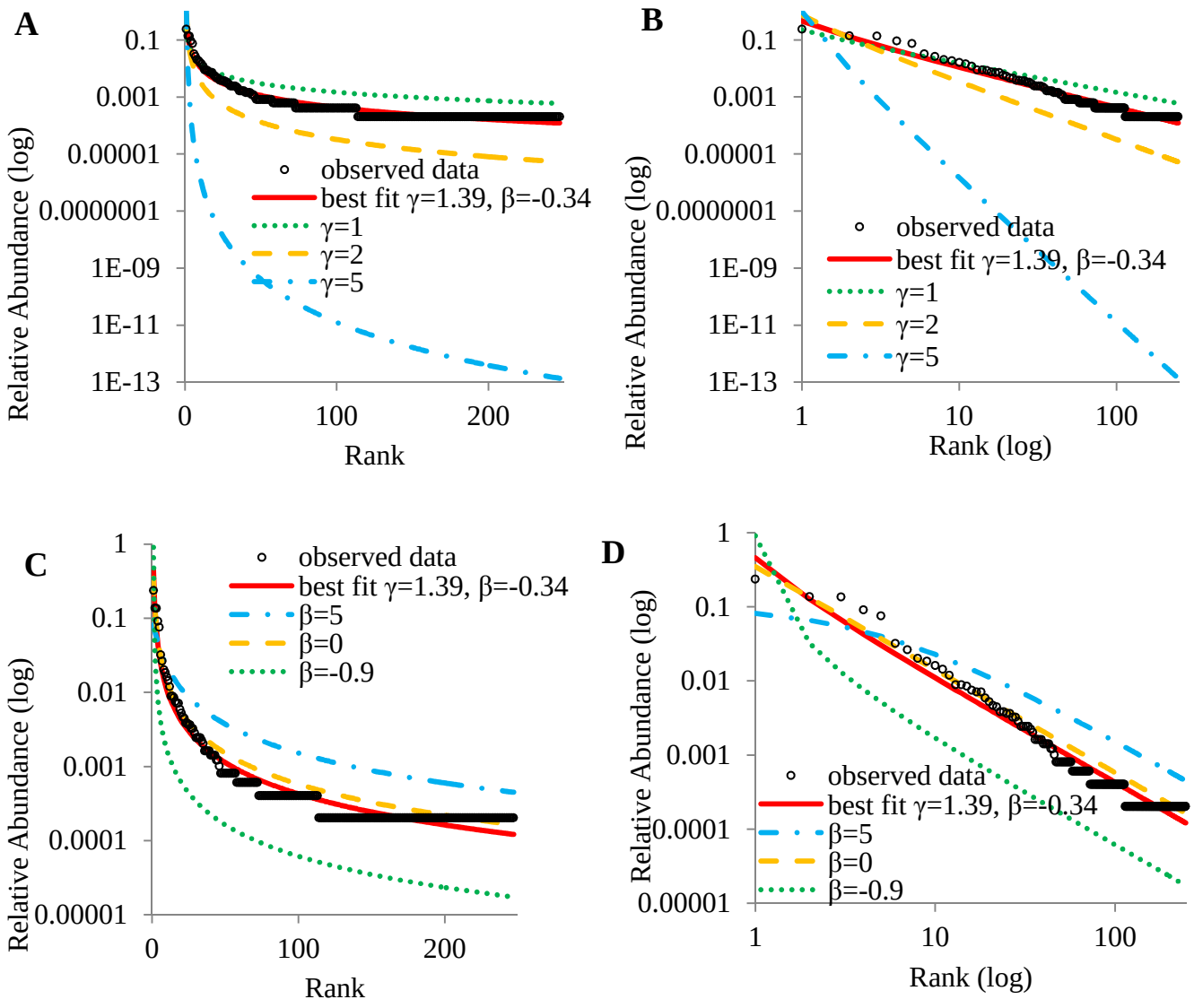

Figure 7.3 Relative abundance distribution (RAD) in case of Zipf-Mandelbrot (ZM) model at different γ (A and B) and β (C and D) values. *S* is 247, sum of p_i was constrained to a value of 1 by varying p_0.

To assess the effect of model parameters on the community diversity, the total species number and individual number were fixed the same as in 1F. As shown in Table 7.2, the increase in γ value results in the decrease in diversity and evenness. However the increase in β value results in the increase in diversity and evenness. In our study the bulk water sample 5P had the largest γ value and smallest β value, therefore, the observed diversity and evenness were the smallest among all the bulk water samples. The sample 5P was separated from other samples because of the large γ value when we pool the bulk water RADs in one figure (Figure 7.4 A and 7.4 B). The most interesting observation in this study is that biofilm developed on the same pipe material had similar γ values and β values, suggesting that the pipe material may play an important role in shaping the bacterial community. The biofilm on PVC pipe had largest β values and copper biofilm had the smallest β values, suggesting that the abundant species on PVC surface were more diverse and more evenly distributed than on the surface of copper pipe. However PVC biofilm also had the largest γ values and copper biofilm γ values were the lowest. The large γ values for PVC biofilm were supposed to level off the total diversity. In this case the β value weighted more than γ value on the diversity, therefore the observed diversity and evenness of PVC biofilm were larger than copper biofilm. The γ and β for galvanized iron biofilm were between the PVC and copper biofilm, so did the bacterial community diversity and evenness. All the biofilm RADs were twisted together because of the negative correlation between γ and β ($R^2 = 0.91$) (Figure 7.4 C and 7.4 D). The negative correlations between the two parameters were also observed by other ecologists (Izsak 2006, Mouillot and Lepretre 2000). The negative relationship between the two parameters was considered caused by opposite affection of environment maturation process on the two parameters: the increase of environmental maturity increases the β value (the evenness of abundant species) and decreases

the γ value (the average possibility of appearance) (Izsak 2006, Wilson et al. 1996). However most evidences were found in the communities in macro ecology the applications of ZM model and the descriptions on succession process indicated by the ZM model for microbial communities need more experimental evidences and further tests.

Henne et al investigated the bulk water and biofilm in drinking water distribution system using DNA and RNA based SSCP. Their RADs showed exponential distribution (Henne et al. 2012). Biofilm had deep slope and small interception than bulk water according to the exponential regression analysis. When we pool samples together, the biofilm and bulk water didn't separate. The power regression also resulted in similar slope and intercept (Figure 7.4E). Inceoglu found that the soil bacterial community RADs were best fit by power law model with the γ value around 0.70 (Inceoglu et al. 2011). The drinking water and biofilm were supposed to have larger γ because they should have smaller diversity than soil bacterial community. However Inceoglu's RADs were plot based on genera level classification. The comparison of the absolute values with literature for the γ and β were not very meaningful because the OTU cluster criteria, the model fitting method, sampling depth and scale were not set in the same level.

Table 7.2 The effects of γ and β on Shannon diversity and evenness

β	γ	Shannon	Evenness
-0.34	1.39	2.58	0.47
-0.34	1	4.05	0.73
-0.34	2	1.15	0.21
-0.34	5	0.06	0.01
-0.9	1.39	0.63	0.11
0	1.39	3.04	0.55
5	1.39	4.41	0.8

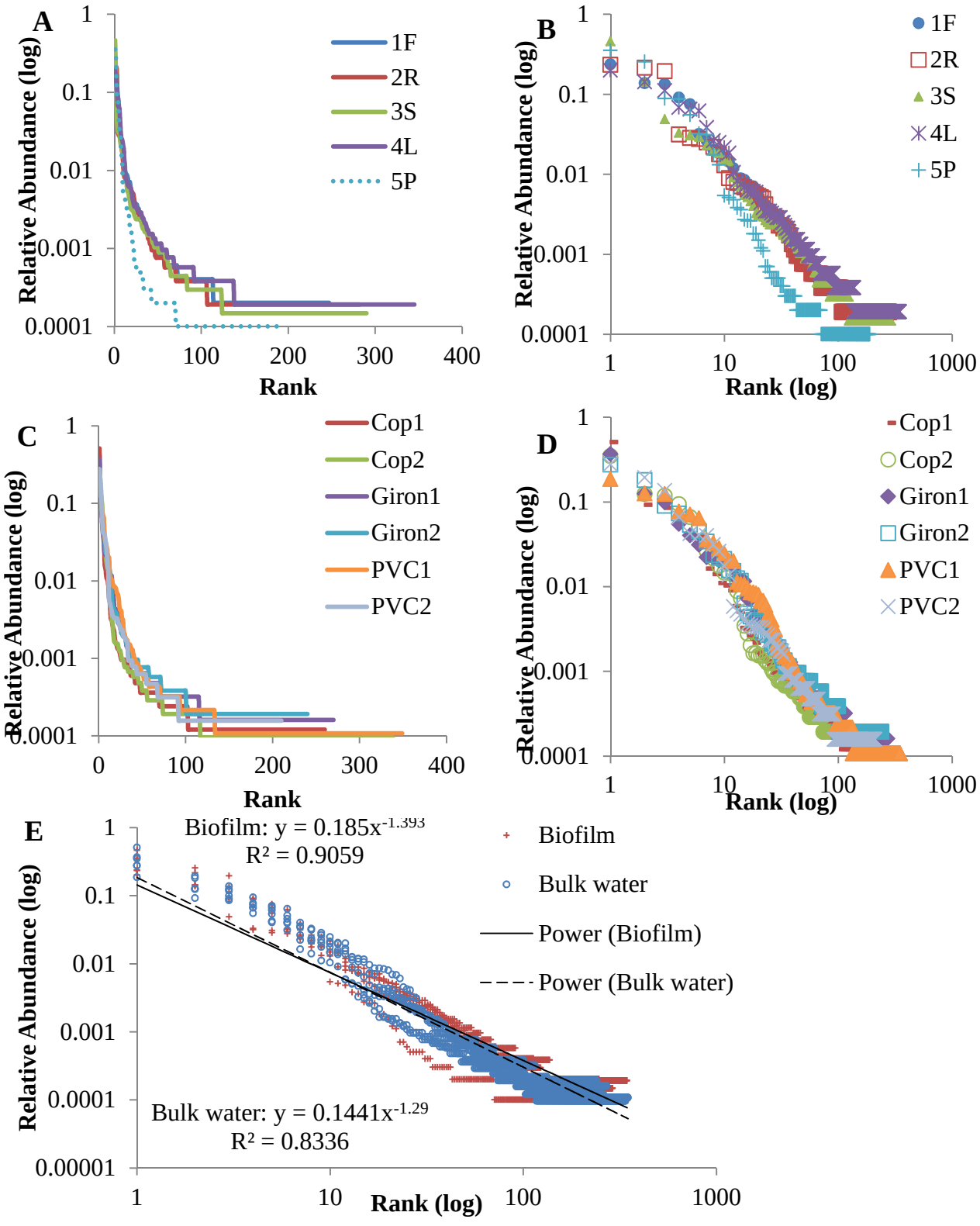

Figure 7.4 Relative abundance distribution (RAD) for bacterial community in drinking water and biofilm.

Using the best fit model for the distribution curve to explain the community assembly mechanism is one possible way. However, some patterns can be described by more than one model. Sometimes either pure statistical model have better fit than models developed from theoretical ecological processes or it's hard to find one best fit ecological model. The link between RAD curve and assembly process may not always act as one to one correspondence. In reality, more than one process may be involved in the assembly of one single community. Biological interactions among species as well as environmental selection may work simultaneously or sequentially in shaping the whole community. The observed community structure occurred as a balance of niche selection, species adaptation, competition, evolution, dispersion and etc. Therefore the conclusion based on RAD and model fitting is still debatable. And statistic description is more acceptable than the ecological explanation.

7.5. Conclusions

Zipf-Mandelbrot model fits all the relative abundance distribution of bacteria communities in both bulk water in drinking water distribution system and biofilm developed inside lab scale pipe lines. ZM model explained about 90% of the total variances for both bulk water and biofilm bacteria community, indicating a community with a few very abundant species coexisting with many rare species. The low β value indicated low niche diversity in drinking water and the surface of pipes. Biofilm developed on the same pipe materials had similar γ and β values suggesting that pipe material may play an important role in the assembly of biofilm bacterial community.

7.6. References

Almaas, E., Kovacs, B., Vicsek, T., Oltvai, Z.N. and Barabasi, A.L. (2004) Global organization of metabolic fluxes in the bacterium Escherichia coli. Nature 427(6977), 839-843.

Craun, G.F., Brunkard, J.M., Yoder, J.S., Roberts, V.A., Carpenter, J., Wade, T., Calderon, R.L., Roberts, J.M., Beach, M.J. and Roy, S.L. (2010) Causes of Outbreaks Associated with Drinking Water in the United States from 1971 to 2006. Clinical Microbiology Reviews 23(3), 507-528.

Dunbar, J., Barns, S.M., Ticknor, L.O. and Kuske, C.R. (2002) Empirical and theoretical bacterial diversity in four Arizona soils. Applied and Environmental Microbiology 68(6), 3035-3045.

Frontier, S. (1985) Diversity and structure in aquatic ecosystems. Oceanography and Marine Biology 23, 253-312.

Galand, P.E., Casamayor, E.O., Kirchman, D.L. and Lovejoy, C. (2009) Ecology of the rare microbial biosphere of the Arctic Ocean. Proceedings of the National Academy of Sciences of the United States of America 106(52), 22427-22432.

Hammes, F., Berney, M., Wang, Y.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 42(1-2), 269-277.

Henne, K., Kahlisch, L., Brettar, I. and Hofle, 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.

Horner-Devine, M.C., Lage, M., Hughes, J.B. and Bohannan, B.J.M. (2004) A taxa-area relationship for bacteria. Nature 432(7018), 750-753.

Hubbell, S.P. (2001) The unified neutral theory of biodiversity and biogeography, Princeton University Press, Princeton.

Huse, S.M., Dethlefsen, L., Huber, J.A., Welch, D.M., Relman, D.A. and Sogin, M.L. (2008) Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. Plos Genetics 4(11).

Inceoglu, O., Abu Al-Soud, W., Salles, J.F., Semenov, A.V. and van Elsas, J.D. (2011) Comparative Analysis of Bacterial Communities in a Potato Field as Determined by Pyrosequencing. Plos One 6(8).

Izsak, J. (2006) Some practical aspects of fitting and testing the Zipf-Mandelbrot model - A short essay. Scientometrics 67(1), 107-120.

Kuznetsov, V.A. (2003). Zhang, W. and Shmulevich, I. (eds), pp. 125-171, Springer US.

Magurran, A.E. (2004) Measuring Biological Diversity, Blackwell, Oxford, UK.

McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., Dornelas, M., Enquist, B.J., Green, J.L., He, F.L., Hurlbert, A.H., Magurran, A.E., Marquet, P.A., Maurer, B.A., Ostling, A., Soykan, C.U., Ugland, K.I. and White, E.P. (2007) Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology Letters 10(10), 995-1015.

Mouillot, D. and Lepretre, A. (2000) Introduction of relative abundance distribution (RAD) indices, estimated from the rank-frequency diagrams (RFD), to assess changes in community diversity. Environmental Monitoring and Assessment 63(2), 279-295.

Narang, R. and Dunbar, J. (2004) Modeling bacterial species abundance from small community surveys. Microbial Ecology 47(4), 396-406.

Nikkari, S., Lopez, F.A., Lepp, P.W., Cieslak, P.R., Ladd-Wilson, S., Passaro, D., Danila, R. and Relman, D.A. (2002) Broad-range bacterial detection and the analysis of unexplained death and critical illness. Emerging Infectious Diseases 8(2), 188-194.

Poitelon, J.B., Joyeux, M., Welte, 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 & Biotechnology 37(2), 117-128.

Polaczyk, A.L., Narayanan, J., Cromeans, T.L., Hahn, D., Roberts, J.M., Amburgey, J.E. and Hill, V.R. (2008) Ultrafiltration-based techniques for rapid and simultaneous concentration of multiple microbe classes from 100-L tap water samples. Journal of Microbiological Methods 73(2), 92-99.

Prosser, J.I., Bohannan, B.J.M., Curtis, T.P., Ellis, R.J., Firestone, M.K., Freckleton, R.P., Green, J.L., Green, L.E., Killham, K., Lennon, J.J., Osborn, A.M., Solan, M., van der Gast, C.J. and Young, J.P.W. (2007) Essay - The role of ecological theory in microbial ecology. Nature Reviews Microbiology 5(5), 384-392.

Schloss, P.D. and Handelsman, J. (2006) Toward a census of bacteria in soil. Plos Computational Biology 2(7), 786-793.

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., Stres, B., Thallinger, G.G., Van Horn, D.J. and Weber, C.F. (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology 75(23), 7537-7541.

Sloan, W.T., Woodcock, S., Lunn, M., Head, I.M. and Curtis, T.P. (2007) Modeling taxa-abundance distributions in microbial communities using environmental sequence data. Microbial Ecology 53(3), 443-455.

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.

Watkins, A.J. and Wilson, J.B. (1994) Plant community structure, and its relation to the vertical complexity of communities: dominance/diversity and spatial rank consistency. Oikos 70(1), 91-98.

Wilson, J.B. (1991) Methods for Fitting Dominance Diversity Curves. Journal of Vegetation Science 2(1), 35-46.

Wilson, J.B., Wells, T.C.E., Trueman, I.C., Jones, G., Atkinson, M.D., Crawley, M.J., Dodd, M.E. and Silvertown, J. (1996) Are there assembly rules for plant species abundance? An investigation in relation to soil resources and successional trends. Journal of Ecology 84(4), 527-538.

Conclusions

Bacterial community composition in two drinking water samples were investigated through16S rRNA gene based pyrosequencing and clone library analysis. Pyrosequencing significantly improved the sampling coverage and revealed a broad diversity of bacterial communities in drinking water, which were dominated by *Alphaproteobacteria* and *Betaproteobacteria*. The bacterial community in drinking water also experienced significant seasonal changes, with *Oxalobacteraceae* succeeding *Methylobacteriaceae* as the predominant bacteria family from winter to summer. These results were consistent with those from a 16S rRNA gene clone library analysis conducted in parallel with pyrosequencing. Phylogenetic analysis further revealed that abundant bacterial populations in drinking water were closely related to metabolically versatile bacterial species broadly distributed in aquatic environments, suggesting a potential link between environmental distribution, metabolic trait, and presence in drinking water.

A small survey performed by investigating bacterial community composition in five geographically distributed drinking water samples. Pyrosequnencing revealed a broad range of diverse bacteria predominated by *Alphaproteobacteria*, *Betaproteobacteria* and *Actinobacteria*. Clone library analysis confirmed the dominant populations detected by pyroseqeuncing. Several core abundant families were detected in all the water samples: *Sphingomonadaceae*, *Caulobacteraceae*, *Methylobacteriaceae*, *Oxalobacteraceae*, *Comamonadaceae*, *Mycobacteriaceae* and *Peptostreptococcaceae*, which represented by *Sphingomonas*, *Novosphingobium*, *Caulobacter*, *Methylobacterium*, *Massilia*, *Acidovorax*, and *Mycobacterium*.

Principal component analysis and cluster analysis showed that bacterial community compositions were influenced by source water and environmental variables

A membrane filtration plant and a conventional sand filtration plant were monitored from raw water to customer's tap water through filtration, chlorination, distribution and stagnation. Bacterial community dynamics during drinking water treatment and distribution processes were investigated by pyrosequencing of the 16S rRNA amplicons. Substantial differences were observed after each treatment and distribution steps for both treatment plants. Membrane filtration removed a large variety of bacteria populations; however, was less effective in removing bacteria from the genus of *Delftia* and *Pseudomonas*. Chlorine disinfection was the key step for bacteria removal, subsequently played an important role in shaping bacteria community structure in tap water. Conventional sand filtration and disinfection also greatly affected the bacterial community composition in water. The distribution had greater influence on the fresh tap water than disinfection. Stagnation showed substantial influence on the bacterial community composition for both water treatment plants. After stagnation, *Betaproteobacteria* greatly decreased, instead *Alphaproteobacteria* and *Firmicutes* greatly increased and became the dominant population. The core resilient populations that survive treatment and distribution process from the family of *Sphingomonadaceae, Burkholderiaceae, Comamonadaceae, Oxalobacteraceae, Xanthomonadaceae*. The occurrence of bacteria members from the genus of *Methylobacterium, Sphingomonas*, *Paracoccus* and *Mycobacterium* in the fresh tap water and stagnated water confirmed isolations indicated potential risk for human health cause by distribution and stagnation.

The influence of source water, disinfection and filtration technology on drinking water bacterial community were investigated by comparing four drinking water treatment plants from the raw river to customer's tap. Bacterial community dynamics during drinking water treatment and distribution processes were investigated through 16S rDNA based pyrosequencing analysis. Chlorination was the key step controlling the bacterial community structure in tap water. Membrane filtration had better treatment efficiency than sand filtration, however, didn't cause significantly different bacterial communities. The influence order on the water bacterial community were Disinfection> Water sources > Filtration. *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, and *Firmicutes* were found dominated in all the water treatment plants. The persistent core bacteria population which survived each treatment and distribution steps and appeared in all the four water treatment plants was from the genus of *Sphingomonas*. The core bacterial populations observed in all the finished water and tap water samples were belonged to the family of *Sphingomonadaceae*, *Comamonadaceae*, *Moraxellaceae* and *Pseudomonadaceae*, which were classified to the bacteria genera of *Sphingomonas*, *Acinetobacter* and *Pseudomonas*.

The effect of pipe materials on biofilm bacterial community composition were investigated through 16S rDNA based pyrosequencing analysis. Biofilm on the surface of copper pipe showed different bacterial community than PVC and galvanized iron. Biofilm developed on the surface of copper pipes was dominated by *Alphaproteobacteria*, characterized by high abundance of *Methylobacteriaceae*, *Erythrobacteraceae*. Biofilm on the surface of PVC and galvanized iron were dominated by both *Alphaproteobacteria* and *Betaproteobacteria*. Half of bacterial populations in each pipe biofilm were also found in other pipes. The core abundant

bacteria families found existed in biofilm of all pipe material surfaces were *Mycobacteriaceae*, *Caulobacteraceae*, *Bradyrhizobiaceae*, *Methylobacteriaceae*, *Erythrobacteraceae*, *Sphingomonadaceae*, *Burkholderiales*_incertae_sedis, *Comamonadaceae* and *Oxalobacteraceae*.

Relative abundance distribution (RAD) of drinking water and biofilm developed in lab scale pipe lines were characterized using Zipf-Mandelbrot (ZM). About 90% of the total variances for both bulk water and biofilm bacteria community structure were explained by ZM model, which indicated that the bacterial communities were all characterized by a few very abundant species followed by a long tail of rare species. Low β values of the bacteria RADs revealed low niche diversity in drinking water and pipe surfaces. Biofilm developed on the same pipe materials had similar RAD shapes as reflected by the similarity of both γ and β values, suggesting that pipe material may play an important role in the biofilm bacterial community assembly.

Vita

Yan Zhang was born in HeBei Province, China, on November 25, 1979. She attended undergraduate school at Shenyang Pharmaceutical University in Shenyang, China and received Bachelor's degree of Microbiological Pharmaceutics in July 2001. She went to graduate school at Shenyang Pharmaceutical University, graduating with a Master's degree of Microbiology and Biochemical Pharmacy in July 2004. Then she admitted as a Ph.D. candidate by Institute of Applied Ecology, Chinese Academy of Sciences, working as a research assistant and received her Ph. D. of Microbiology in July 2007. She went to the University of Tennessee, the United States in January, 2008 and received her second Doctorate in Civil Engineering in December, 2012.