# Effect of Water Chemistry, Pipe Material, Temperature and Flow on the Building Plumbing Microbiome and Opportunistic Pathogen Occurrence

Pan Ji

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Civil Engineering

Amy J. Pruden-Bagchi, Chair
Marc A. Edwards
Brian D. Badgley
Peter J. Vikesland

September 25, 2017
Blacksburg, VA

Keywords: Building plumbing microbiome, opportunistic pathogens, water chemistry, temperature

Effect of Water Chemistry, Pipe Material, Temperature and Flow on the Building Plumbing Microbiome and Opportunistic Pathogen Occurrence

Pan Ji

## ABSTRACT

The building plumbing microbiome has important implications, especially in terms of its role as a reservoir and conduit for the spread of opportunistic pathogens (OPs), such as *Legionella pneumophila*. This dissertation applied next-generation DNA sequencing tools to survey the composition of building plumbing microbiomes and assessed hypothetical factors shaping them.

A challenge to identifying key factors shaping building plumbing microbiomes is untangling the relative contributions of influent water quality, provided by drinking water utilities, and those of building-level features, such as pipe materials. To this end, standardized pipe rigs were deployed at the treatment plants and in distal portions of the water distribution system at five water utilities across the eastern U.S. Source water and treatment practices appeared to be the overarching factors shaping the microbial taxonomic composition at the tap, with five key water chemistry parameters identified (total chlorine, pH, P, $SO_4^{2-}$ and $Mg^{2+}$).

Hot water plumbing is of particular interest because OPs tend to proliferate in warm water environments and can be inhaled in aerosols when showering. Two identical lab-scale recirculating hot water rigs were operated in parallel to examine the combined effects of water heater temperature set point, pipe orientation, and water use frequency on the hot water plumbing microbiome. Our results revealed distinct microbial taxonomic compositions between the biofilm and water phases. Importantly, above a threshold of 51°C, water heater temperature, pipe orientation, and water use frequency together incurred a prominent shift in microbiome composition and *L. pneumophila* occurrence.

While heat shock is a popular means of remediating *L. pneumophila* contamination in plumbing, its broader effects on the microbiome are unknown. Here, heat shock was applied to acclimated

lab-scale hot water rigs. Comparison of pre- versus post- heat shock samples indicated little to no change in either the microbial composition or *L. pneumophila* levels at the tap, where both water heater temperature and water use frequency had the most dominant effect.

Overall, this dissertation contributes to advancing guidance regarding where to most effectively target controls for OPs and also advances research towards identifying the features of a "healthy" built environment microbiome.

Effect of Water Chemistry, Pipe Material, Temperature and Flow on the Building Plumbing Microbiome and Opportunistic Pathogen Occurrence

Pan Ji

## GENERAL AUDIENCE ABSTRACT

Drinking water is often misconceived to be "sterile," whereas in reality the water distribution and plumbing systems that convey the water to the consumer represent a robust microbial habitat. While it is not possible, or even desirable, to kill all of the microbes present in drinking water, the Safe Drinking Water Act in the U.S. enforces measures to purify and disinfect water at the treatment plant and keep bacterial numbers low in water mains and up to the consumer property line. However, current regulatory frameworks are designed to protect against fecal- (e.g., raw sewage and manure) derived pathogens, whereas recently opportunistic pathogens (OPs), including *Legionella pneumophila*, *Mycobacterium avium*, and *Pseudomonas aeruginosa* have come to the forefront as the leading source of tap-water related illness in the U.S. and other developed countries. In contrast to traditional fecal pathogens, building plumbing systems are a natural habitat for OPs, where they can readily proliferate. Currently there are no provisions within the Safe Drinking Water Act or other regulations to protect consumers specifically from OPs. There are also no "silver bullet" remedial measures that consistently and reliably defend against OPs colonizing building building plumbing, particularly when aiming to protect against multiple types of OPs. A major challenge in preventing and remediating OP proliferation in building plumbing is that they tend to be protected from disinfectants, such as chlorine, inside amoeba hosts and within the slimy layer that forms on the surface of pipe walls called "biofilm".

With the recent advent over the past decade of next-generation DNA sequencing, there are new reasons to take interest in the microbial composition of tap water. In particular, next-generation DNA sequencing has provided new insight into the composition of the human microbiome, e.g., the microbes that naturally inhabit our skin, gut, and lungs, and has revealed striking relationships with human health (e.g., obesity, diabetes, asthma, autism, allergies). The question naturally arises with respect to the factors shaping the human microbiome, with role of the "built

environment" being of fundamental interest. The built environment; including homes, offices, schools, hospitals, and vehicles, is where most humans in developed countries spend > 90% of their time. Tap water is likely an important feature shaping the microbiome of the built environment, serving as a conduit for microbes into tiny droplets called aerosols, which can be inhaled into the lungs or otherwise inoculate the skin during showering or be transferred onto food during food preparation. Thus, there is interest in mapping out the microbiome of tap water and the factors that shape it, not only because of its potential to harbor OPs, but because of its potential general effect on built environment and human microbiomes. Long-term research could lead towards identifying which microbes serve a beneficial, or "probiotic," role in preventing pathogen growth and benefiting human health.

The purpose behind the body of research described in this dissertation was to apply newly available next-generation DNA sequencing tools towards mapping out the microbial composition characteristic of tap water, with emphasis on implications for preventing proliferation of OPs. Of particular interest was the relative role of what water utilities and building operators can do to protect public health. To this end, the DNA sequencing approach was applied to carefully controlled and replicated field- and laboratory-scale plumbing rigs to gain insight into the relative roles and interactions of the water quality provided by drinking water utilities and practical building-level engineering controls. Specific factors investigated included: stagnation (i.e., the tendency of water to sit unused in pipes in 8 hour cycles), pipe material (e.g., metallic versus plastic), pipe configuration (i.e., up or down flow to induce convective mixing vs stratification, respectively), water heater temperature set point (i.e., balancing hotter temperatures needed to kill pathogens versus lower temperatures desirable to save energy or prevent scalding), and heat-shock treatment (i.e., temporarily elevating the water heater temperature and flushing the system to kill off pathogens).

There were several general findings that can be highlighted based on this research. First, based on comparison of standardized plumbing rigs installed at five water utilities in the U.S., the nature of the water provided by the local water utility was the overarching factor shaping the microbiome composition at the tap, moreso than pipe material or stagnation. Second, there exists an ideal threshold water heater temperature setting (51 °C based on the conditions of this

study) above which there is a concordant shift in microbiome composition and decrease in *L. pneumophila* occurrence. Third, consistent water heater temperature setting above this threshold has a stronger long-term influence on the microbiome composition and *L. pneumophila* control than temporarily elevating the temperature for heat-shock treatment. Finally, biofilm and bulk water microbial compositions are extremely diverse in composition (e.g., thousands of species of microbes in each) and functional markers, and distinct from one anaother in terms of their characteristics under different operational conditions.

In sum, this study takes a step towards better understanding building plumbing microbiome and identifies several promising engineering and control factors that can ultimately inform intentional engineering of the building plumbing microbiome, particularly with respect to protecting public health against OPs and potentially other microbiome-related ailments in the future.

## AUTHOR'S PREFACE

The four chapters comprising the main body of this dissertation (Chapters 2 – 5) are arranged as individual manuscripts according to Virginia Tech's specifications and are formatted based on the journal to which it was published or submitted. Each manuscript was produced in collaboration with Dr. Amy Pruden, Dr. Marc Edwards, and other researchers (Dr. Jeffrey Parks and Dr. William J. Rhoads) with intellectual contributions appropriately indicated by authorship or co-authorship. Dr. Pruden and Dr. Edwards are project PI and co-PI, respectively, who both contributed to experiment design, data interpretation, and manuscript review (Chapters 2 – 5). Dr. Jeffrey Parks (Chapter 2) contributed to field rig design, sample acquisition, water chemistry analysis, and manuscript review. Dr. William J. Rhoads (Chapters 3 – 5) contributed to installation and daily-maintanence of lab-scale rig system, sample acquisition, DNA extraction, OP gene copy numbers quantification (Chapter 4), and manuscript review. Throughout the dissertation, the next-generation DNA sequencing tool, Illumina 16S rRNA gene amplicon sequencing, is applied as a relatively new method for profiling microbial community composition

**CHAPTER 1** is a brief overview, which outlines the state of the knowledge with respect to opportunistic pathogen (OP) occurrence and microbiome composition in building plumbing, knowledge gaps, and a roadmap of this dissertation.

**CHAPTER 2** examines the interplay among water chemistry, pipe material, typical 8-hour diurnal stagnation cycles and water microbiome composition taking advantage of field-scale standardized and replicated pipe rigs that were already installed and acclimated prior to this study at five water utilities across U.S. Chapter 2 is the only chapter in this dissertation focused specifically on cold tap water, with the remaining examining hot water plumbing sytems. Chapter 2 has been published in *PLoS One*: Ji P, Parks J, Edwards MA, Pruden A. (2015). Impact of water chemistry, pipe material and stagnation on the building plumbing microbiome. *PLoS ONE*, 10(10): e0141087.

**CHAPTER 3** investigates the impact of water heater temperature setting and water use frequency on hot water plumbing microbiome through implementing parallel lab-scale hot water systems. The purpose of the study is to understand how design and operation of hot water plumbing systems can alter the microbiome in the hot water plumbing, which has broader implications on energy conservation and public health. Chapter 3 has been published in *ISME J*: Ji P, Rhoads WJ, Edwards MA, Pruden A. (2017). Impact of water heater temperature setting and water use frequency on the building plumbing microbiome. *ISME J*. 11: 1318-1330.

**CHAPTER 4** extended the research in **Chapter 3** by exploring the effect of heat shock as the most widely adopted remedial or management measure to address concerns about OP growth in hot water plumbing. The limited influence of heat shock in shaping the broader plumbing microbial community composition corroborated findings in Chapter 3 that maintaining a sufficiently high water heater temperature setting is the most effective way to control *Legionella* and, correspondingly, a sustained elevated temperature has a stronger effect on the microbial community composition than short term heat shock. Chapter 4 has been submitted to *Microbiome*: Ji P, Rhoads WJ, Edwards MA, Pruden A. (2017). Effect of heat shock on hot water plumbing microbiota and *Legionella pneumophila* control. *Microbiome*. Revision Submitted.

**CHAPTER 5** probes functional features of microbial communities inhabiting hot water plumbing, especially as they correspond to various engineering control interventions, using more recently available shotgun metagenomic DNA sequencing analysis. Moving into the future, such metagenomics methods are hoped to go beyond taxonomic identification of the kinds of microbes present and to provide deeper insight into what functions microbial communities are actually performing. Specifically, this study aimed to further characterize associations between opportunistic pathogens and the broader microbial community. Chapter 5 summarizes preliminary data analysis results, with the intention to submit the manuscript to a journal for publication.

**CHAPTER 6** summarizes the major findings of this work and discusses future work.

In addition to the 4 manuscripts included in this dissertation, the entire PhD research work included several other collaborative manuscripts not included in this dissertation research as follows:

Published:

1. Muñoz-Egea MC, **Ji P**, Pruden A, Falkinham III JO. (2017). Inhibition of adherence of *Mycobacterium avium* to plumbing surface biofilms of *Methylobacterium* spp. *Pathogen* 6(3): 42. doi:10.3390/pathogens6030042.
2. Rhoads WJ, Garner E, **Ji P**, Zhu N, Parks J, Schwake D, Pruden A, Edwards M. (2017). Distribution system operational deficiencies coincide with reported Legionnaires' disease clusters in Flint, MI. *Environ Sci Technol*. Accepted. doi: 10.1021/acs.est.7b01589.
3. Buse HY, **Ji P**, Gomez-Alverez V, Pruden A, Edwards MA, Ashbolt NJ. (2016). Effect of temperature and colonization of *Legionella pneumophila* and *Vermamoeba vermiformis* on bacterial community composition of copper drinking water biofilms. *Microb Biotechnol* 10(4): 773 – 788. doi: 10.1111/1751-7915.12457.
4. Rhoads WJ, **Ji P**, Pruden A, Edwards MA. (2015). Water heater temperature set point and water use patterns influence *Legionella pneumophila* and associated microorganisms at the tap. *Microbiome* 3: 67. doi: 10.1186/s40168-015-0134-1.

In preparation:

1. Ji P, Rhoads WJ, Edwards MA, Pruden A. (2017). Effect of cupric ion and antimicrobial on hot water plumbing microbiome.

## ACKNOWLEDGEMENTS

If all acknowledgements were to be duly noted, this section would be at least the same length of the entire dissertation. To protect the author from becoming too emotional to proceed with the dissertation writing, below are the most important personnel and groups that have significantly contributed to the generation of this dissertation.

Drs. Amy Pruden and Marc Edwards are not only the author's dearest co-advisors, but also amazing researchers, insightful mentors, and great people. There are so many things the author wish to thank them for, large and small. With them, the PhD adventure became an enjoyable journey, and the author could not have asked for more. As committee members, Dr. Peter Vikesland and Dr. Brian Badgley offered valuable input to this dissertation, and generous help whenever being requested.

The Pruden and Edwards lab members accompanied the author during the past five years and shared numerous happy memories. The author feel privileged to work with such a brilliant and fun group of people. Several online communities have consistently enlightened and inspired the author on programing, data mining and visualization, including Stack Overflow and R-bloggers.

Finally, the author would like to thank the family and friends for their unconditional support.

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# CHAPTER 1. OVERVIEW

It is not until recently that the built environment microbiome, i.e., the collection of microbes that inhabit the built environment, have been credited for their influence on the health of modern humans (Hoisington et al., 2015). For example, the microbiome of humans living in the jungles of Amazonia are distinct from those of individuals living in highly urbanized areas (Yatsunenko et al., 2012; Blaster et al., 2013). Given that modern humans spend upwards of 95% of their time indoors, it is critical to understand: 1) what is the composition of the indoor microbiome? 2) what are the factors shaping the indoor microbiome? And 3) what are the implications for human health? To this end, The Alfred P. Sloan Foundation recently launched a new interdisciplinary field of inquiry dubbed "Microbiology of the Built Environment (MoBE)" (www.microbe.net). The fields represented by the consortium of researchers brought together by this initiative included microbiology, microbial ecology, water chemistry, air quality, building science, and Architecture, with the advent of next-generation DNA sequencing being a major driver in the powerful new capability that it brought in quickly and economically being able to probe the genes of the tens to hundreds of thousands of different microbes characteristic of the air, water, and surfaces encountered in the built environment (Wilkins et al., 2017). Advances in next-generation DNA sequencing techniques have accelerated endeavors to map the "uncultivated majority" (Rappé and Giovannoni, 2003) and has laid the cornerstone for unprecedented scale of microbiome-centered studies, including the Human Microbiome Project (The Human Microbiome Project Consortium, 2012) and the Earth Microbiome Project (Gilbert et al., 2014).

Herein, this dissertation specifically focuses on microbes inhabiting premise (i.e., building) plumbing systems as a key component of the built environment microbiome. Building plumbing is defined as the portion of the distribution system from the water meter to the point-of-use at the tap in residential homes and other buildings (National Research Council, 2005). Building plumbing essentially provides a complex freshwater habitat for a diverse array of microbes, delivering them into the built environment as the tap water is used routinely for food preparation, drinking, showering, and hygiene. Presently, very little is known about who these microbes are and their implications for human health. Historically, most water quality and regulatory attention has been directed to fecal pathogens, such as *Escherichia coli* and *Shigella* (Edberg et

al., 2000). However, fecal pathogens are typically dealt with effectively at the water treatment plant through filtration and disinfection, and do not subsequently survive well as clean, disinfected water proceeds through the municipal water distribution system and into the home plumbing- environments that are distinct from the warm gastrointestinal tracts from which they originated (Falkinham III et al., 2015a). Instead, a new threat has emerged in recent decades, one from microbes that are adaptive to the relatively harsh, oligotrophic environment encountered in drinking water systems. Specifically, opportunistic pathogens (OPs); such as *Legionella pneumophila*, *Mycobacterium avium*, *Pseudomonas aeruginosa*, *Acanthamoeba polyphaga*, and *Naegleria fowleri* have presented themselves as a newer and increasing concern as collectively they have become the primary source of tap water-related illness, outbreak, and death in developed countries (Falkinham III et al., 2015a). OPs especially pose a risk to immunocompromised populations (e.g., immunodeficiency, aging, pressure) (Brown et al., 2012), but the proportion of the population at risk is steadily increasing (Falkinham III et al., 2015b). OPs are also challenging because a primary route of exposure is inhalation of aerosols, rather than the standard paradigm of exposure via ingestion (Falkinham III et al., 2015a).

The general framing and approach of this dissertation is to apply newly available next-generation DNA sequencing techniques to explore the vastly unknown composition of the microbiome of building plumbing, while specifically focusing on implications for OP control within the context of practical engineering design and management practices. Factors investigated include water treatment and distribution characteristics, pipe material, water heater temperature set point, water use frequency, flow configuration, and remedial measures such as "heat shock." OPs are notoriously challenging to address because they are natural inhabitants and often recalcitrant to standard treatment processes, which can actually facilitate their regrowth and enrichment at the tap. OPs share several common traits including heat-tolerance, disinfectant-resistance (Falkinham III et al., 2015a), and ability to survive and proliferate within amoeba hosts (Thomas and Ashbolt, 2011). Biofilm, or the slimy microbial layer forming on the inside of pipes, is a particularly important niche in which OPs establish, grow, and can be protected from disinfectants and other assaults (Falkinham III, 2015). In particular, the building plumbing environment has a high surface to volume ratio and typically involves long stagnation times at distal taps near the point of use, which together enhances OP regrowth by supporting biofilm

formation. An important aspect of the work described in this dissertation is systematic comparison of the microbial community and OP composition in both the biofilm and bulk water together through various engineering and design control scenarios in seeking to understand their interrelationship.

Of particular interest to this effort is moving towards distinguishing "good" microbes from the "bad". Arguably, in the U.S. and other developed countries, there has been overemphasis on indiscriminantly killing microbes, through heavy reliance on disinfectants such as chlorine, which is thought to play a role in selecting for OPs (Falkinham III, 2015). Given that it is impossible and even undesirable to achieve sterile tap water, a more sensibal approach may be to think more in terms of "managing" rather than "eradicating" the microbial communities inhabiting building plumbing. This has recently been termed a "probiotic" approach, in which, ideally, conditions can be identified that select for harmless microbes that outcompete OPs in the building plumbing environment (Berry et al., 2006; Wang et al., 2013). While "probiotic" directly implies addition of beneficial microbes, an analogy more likely to play out in the drinking water landscape is a "pre-biotic" approach (de Vrese and Schrezenmeir, 2008; Gourbeyre et al., 2009; Patel and DuPont et al., 2015), in which the conditions in the building plumbing (temperature, disinfectants, nutrients, flow, etc.) are specifically tailored to select for the most benign microbes. This could potentially shift the traditional waterborne-pathogen-control paradigm into an era of microbiome-level control for drinking water treatment and distribution process. It is thus imperative to form fundamental knowledge on the association between OPs and other microbes, as well as the partition/interaction between biofilm and bulk water phases (Declerck, 2010). Next-generation DNA sequencing technology allows for probing microbiome composition at unprecedented depth with minimal cost (Erlich, 2015), providing new technology to support the goal of a probiotic approach.

From an engineering point of view, field application of the probiotic approach involves selecting the engineering design and operation conditions that support a healthy microbiome (*e.g.,* diminished occurrence of OPs). For example, pipe material influences the building plumbing microbiome via multiple routes. Pipe materials differ in their biofilm formation potential (Kerr et al., 1998; van der Kooij et al., 2005; Yu et al., 2010), disinfectant wall-decay rate (Hallam et

al., 2002, free chlorine; Mutoti et al., 2007, combined chlorine), and oxygen permeability (Whelton et al., 2010). Metal pipes release corresponding metal ions, while plastic pipes release an array of chemical compounds which potentially serve as carbon source (Rożej et al., 2015). In particular, copper pipe is known for its antimicrobial effect, *e.g.,* "contact killing" of copper surface (Grass et al., 2011).

Water chemistry is clearly of central importance in governing the water microbiome composition and occurrence of OPs. Different disinfectant type (i.e., chlorine versus chloramine) can shape distinct drinking water microbiome (Williams et al., 2005; Gormez-Alvarez et al., 2012). Particularly in the absence of disinfectant or once disinfectant has decayed, assimilable organic carbon (AOC) is widely viewed as a predictor of heterotrophic bacterial growth potential in drinking water (Douterelo et al., 2014). Both the concentration and composition of AOC can influence pathogens in drinking water (Vital et al., 2010), with varying impact on different OPs (van der Wielen and van der Kooij, 2012). Presence of Fe (> 0.095 ppm) can enhance *L. pneumophila* colonization in hot water recirculation system while Cu (0.76 ppm) diminishes colonization (Serrano-Suárez et al., 2013). Still, there is little understanding of which aspects of water quality are most critical in shaping the water microbiome. Moreover, the relative importance of water chemistry to other engineering conditions remains largely unexplored. Water heater temperature setting has been reported as a crucial factor for OPs control. In a pilot study in the U.S. and Canada, households with water heater temperatures below 50 °C had been associated with higher frequency of detection for nontuberculous mycobacteria (NTM) compared to households with water heater temperatures above 55 °C (Falkinham III, 2011). Nevertheless, OPs can dynamically adapt to high temperature. In a 14-month survey of the drinking water supply system at Braunschweig, Germany, a thermophilic *Legionella* community (including *L. pneumophila*) grew in hot water above 50 °C (Lesnik et al., 2016). A third parameter to consider is the water flow pattern/ water use frequency, which determines the duration of stagnation and further introduces variation in water microbial quality at taps within the same building (Suchomel et al., 2013). Previous reports from our research group has focused on OPs response to pipe material (PVC, cement, iron), disinfectant type (chlorine or chloramine) and water age (1 – 5.7 days) (Wang et al., 2014), while this project will broaden the scope to the microbiome level and specifically emphasize interactions among microbes.

To address the key knowledge gaps and strive towards the overarching goal of this effort as summarized above, four key questions are addressed in four corresponding chapters comprising the body of this disseration. The ultimate findings are hoped to facilitate decision making by building owners and other stakeholders to control OP occurrence at the tap and work towards a new paradigm of probiotic control of the tap-water associated microbiome.

**Summary of Chapters 2 – 5**:

***Q1. To what extent can building owners control their plumbing microbiome (relative control and responsibility between water utilities and building owners)?*** This question involves comparing the relative importance of properties of influent water to buildings versus building-specific plumbing configuration and water use patterns. This is achieved by considering finished water as it exits the treatment plant versus water that has aged after passing through the distribution system and comparing outcomes when fed to identical simulated building plumbing.

***Q2. To what extent can operating conditions shift the building plumbing microbiome?*** Once plumbing systems are emplaced, in reality only a handful of factors can be readily altered by building owners without reconstruction (e.g., water heater temperature setting, flushing tap before usage, etc.). It is critical to understand if and to what extent the building plumbing microbiome shifts in response to these practical shifts in operating conditions.

***Q3. How resilient is the plumbing microbiome in the face of short-term versus long-term shifts in operating conditions are imposed?*** For instance, electricity outages or seasonal building occupancies may result in temporarily switching heaters off, boosting potential pathogen growth at lower temperature. On the other hand, building owners commonly elevate temperatures temporarily as a thermal disinfection (heat treatment) measure.

***Q4. From a microbial ecology point of view, which microbes are most closely associated (positively or negatively) with opportunistic pathogens (OPs)?*** The vast majority of the studies have focused on of OPs in isolation, or at a minimum in relationship with host amoebae

(Swanson and Isberg, 1995; Wang et al., 2012, Lu et al., 2015). In contrast, very little is known about other "background" microbes and their potential to support or inhibit OP proliferation. While numerically OPs typically account for a minor portion of the drinking water microbiome, such small amounts can still be sufficient to incur disease (Schoen and Ashbolt, 2011). Building towards an understanding of the microbial ecology of OPs in this larger context is critical to identify new and more broadly effective OP control paradigms.

## REFERENCES


Berry D, Xi C, Raskin L. (2006). Microbial ecology of drinking water distribution systems. *Curr Opin Biotechnol* 17(3): 297 – 232. doi: 10.1016/j.copbio.2006.05.007.

Blaser MJ, Dominguez-Bello MG, Contreras M, Magrris M, Hidalgo G, Estrada I, et al. (2013). Distinct cutaneous bacterial assemblages in a sampling of South American Amerindians and US residents. *ISME J* 7(1): 85 – 95. doi: 10.1038/ismej.2012.81.

Brown SP, Cornforth DM, Mideo N. (2012). Evolution of virulence in opportunistic pathogens: generalism, plasticity, and control. *Trends Microbiol* 20(7): 336 – 342. doi:10.1016/j.tim.2012.04.005.

Declerck P. (2010). Biofilm: the environmental playground of *Legionella pneumophila*. *Environ Microbiol* 12(3): 557 – 566. doi: 10.1111/j.1462-2920.2009.02025.x.

de Vrese M, Schrezenmeir J. (2008). Probiotics, prebiotics, and synbiotics. *Adv Biochem Eng Biotechnol* 111: 1 – 66. doi: 10.1007/10_2008_097.

Douterelo I, Boxall JB, Deines P, Sekar R, Fish KE, Biggs CA. (2014). Methodological approaches for studying the microbial ecology of drinking water distribution systems. *Water Res* 65(15): 134 – 156. doi: 10.1016/j.watres.2014.07.008.

Erlich Y. (2015). A vision for ubiquitous sequencing. *Genome Res* 25(10): 1411 – 1416. doi: 10.1101/gr.191692.115.

Edberg SC, Rice EW, Karlin RJ, Allen MJ. (2000). *Escherichia coli*: the best biological drinking water indicator for public health protection. *J Appl Microbiol* 88(51): 1065 – 1165. doi: 10.1111/j.1365-2627.2000.tb05338.x.

Falkinham III JO. (2011). Nontuberculous mycobacteria from household plumbing of patients with nontuberculous mycobacteria disease. *Emerg Infect Dis* 17: 419 – 424. doi:10.3201/eid1703.101510.

Falkinham III JO. (2015). Common features of Opportunistic Premise Plumbing Pathogens. *Int J Environ Res Public Health* 12(5): 4533 – 4545. doi: 10.3390/ijerph120504533.

Falkinham III JO, Hilborn ED, Arduino MJ, Pruden A, Edwards MA. (2015a). Epidemiology and ecology of opportunistic premise plumbing pathogens: *Legionella pneumophila*, *Mycobacterium avium*, and *Pseudomonas aeruginosa*. *Environ Health Perspect* 123(8): 749 – 758. doi: 10.1289/ehp.1408692.

Falkinham III JO, Pruden A, Edwards M. (2015b). Opportunistic Premise Plumbing Pathogens: increasingly important pathogens in drinking water. *Pathogens* 4(2): 373 – 386. doi: 10.3390/pathogens4020373.

Grass G, Rensing C, Solioz M. (2011). Metallic copper as an antimicrobial surface. *Appl Environ Microbiol* 77(5): 1541 – 1547. doi:10.1128/AEM.02766-10.

Gilbert JA, Jansson JK, Knight R. (2014). The Earth Microbiome project: successes and aspirations. *BMC Biology* 12: 69. doi: 10.1186/s12915-014-0069-1.

Gomez-Alvarez V, Revetta RP, Domingo JWS. (2012). Metagenomic analyses of drinking water receiving different disinfection treatments. *Appl Environ Microbiol* 78(17): 6095 – 6102. doi: 10.1128/AEM.01018-12.

Gourbeyre P, Denery S, Bodinier M. (2011). Probiotics, prebiotics, and synbiotics: impact on the gut immune system and allergic reactions. *J Leukoc Biol* 89(5): 685 – 695. doi: 10.1189/jlb.1109753.

Hallam NB, West JR, Forster CF, Powell JC, Spencer I. (2002). The decay of chlorine associated with the pipe wall in distribution systems. *Water Res* 36(14): 3479 – 3488. DOI: 10.1016/S0043-1354(02)00056-8.

Hoisington AJ, Brenner LA, Kinney KA, Postolache TT, Lowry CA. (2015). The microbiome of the built environment and mental health. Microbiome 3: 60. doi: 10.1186/s40168-015-0127-0.

Kerr CJ, Osborn KS, Robson GD, Handley PS. (1998). The relationship between pipe material and biofilm formation in a laboratory model system. *J Appl Microbiol* 85 (Suppl 1): 29S – 38S.

Lesnik R, Brettar I, Höfle MG. (2016). *Legionella* species diversity and dynamics from surface reservoir to tap water: from cold adaptation to thermophily. *ISME J* 10: 1064 – 1080. doi: 10.1038/ismej.2015.199.

Lu J, Struewing I, Vereen E, Moe C, Ashbolt N. (2015). Molecular detection of *Legionella* spp. and their associations with *Mycobacterium* spp., *Pseudomonas aeruginosa* and amoeba hosts in a drinking water distribution system. *J Appl Microbiol* 120(2): 509 – 521. doi: 10.1111/jam.12996.

Mutoti G, Dietz JD, Arevalo J, Taylor JS. (2007). Combined chlorine dissipation: pipe material, water quality, and hydraulic effects. *J AWWA* 99(10): 96-106.

National Research Council (NRC). (2005). Public water supply distribution systems: assessing and reducing risks – first report. The National Academies Press.

Patel R, DuPont HL. (2015). New approaches for bacteriotherapy: prebiotics, new-generation probiotics, and synbiotics. *Clin Infect Dis* 60(suppl 2): S108 – S121. doi: 10.1093/cid/civ177.

Rappé MS, Giovannoni SJ. (2003). The uncultured microbial majority. *Annu Rev Microbiol* 57: 369 – 394. doi: 10.1146/annurev.micro.57.030502.090759.

Rożej A, Cydzik-Kwiatkowska A, Kowalska B, Kowalski D. (2015). Structure and microbial diversity of biofilms on different pipe materials of a model drinking water distribution systems. *World J Microbiol Biotechnol* 31: 37 – 47. doi:10.1007/s11274-014-1761-6.

Schoen ME, Ashbolt NJ. (2011). An in-premise model for *Legionella* exposure during showering events. *Water Res* 45(18): 5826 – 5826. doi: 10.1016/j.watres.2011.08.031.

Serrano-Suárez A, Dellundé J, Salvadó H, Cervero-Aragó S, Méndez J, Canals O, et al. (2013). Microbial and physicochemical parameters associated with *Legionella* contamination in hot water recirculation systems. *Environ Sci Pollut Res Int* 20(8): 5534 – 5544. doi: 10.1007/s11356-013-1557-5.

Suchomel M, Diab-Elschahawi M, Kundi M, Assadian O. (2013). Influence of pre- and post-usage flushing frequencies on bacterial water quality of non-touch water fittings. *BMC Infect Dis* 13: 402. doi: 10.1186/1471-2334-13-402.

Swanson MS, Isberg RR. (1995). Association of *Legionella pneumophila* with the macrophage endoplasmic reticulum. *Infect Immun* 63(9): 3609 – 3620.

The Human Microbiome Project Consortium. (2012). A framework for human microbiome research. *Nature* 486: 215 – 221. doi: 10.1038/nature11209.

Thomas JM, Ashbolt NJ. (2011). Do free-living amoebae in treated drinking water systems present an emerging health risk? *Environ Sci Technol* 45(3): 860 – 869. doi: 10.1021/es102876y.

van der Kooij D, VeenendaalHR, Scheffer WJ. (2005). Biofilm formation and multiplication of *Legionella* in a model warm water system with pipes of copper, stainless steel and cross-linked polyethylene. *Water Res* 39(13): 2789 – 2798. doi: 10.1016/j.watres.2005.04.075.

van der Wielen PW, van der Kooij D. (2013). Nontuberculous mycobacteria, fungi, and opportunistic pathogens in unchlorinated drinking water in the Netherlands. *Appl Environ Microbiol* 79(3): 825 – 834. doi: 10.1128/AEM.02748-12.

Vital M, Stucki D, Egli T, Hammes F. (2010). Evaluating the growth potential of pathogenic bacteria in water. *Appl Environ Microbiol* 76(19): 6477 – 6484. doi: 10.1128/AEM.00794-10.

Wang H, Edwards M, Falkinham III JO, Pruden A. (2012). Molecular survey of the occurrence of *Legionella* spp., *Mycobacterium* spp., *Pseudomonas aeruginosa*, and amoeba hosts in two chlorinated drinking water distribution systems. *Appl Environ Microbiol* 78(17): 6285 – 6294. doi: 10.1128/AEM.01492-12.

Wang H, Edwards MA, Falkinham III JO, Pruden A. (2013). Probiotic approach to pathogen control in premise plumbing systems? A review. *Environ Sci Technol* 47: 10117 − 10128. doi: 10.1021/es402455r.

Wang H, Master S, Edwards MA, Falkinham III JO, Pruden A. (2014). Effect of disinfectant, water age, and pipe materials on bacterial and eukaryotic community structure in drinking water biofilm. *Environ Sci Technol* 48(3): 1426 – 1435. doi: 10.1021/es402636u.

Whelton A, Dietrich A, Gallagher D. (2010). Contaminant diffusion, solubility, and material property differences between HDPE and PEX potable water pipes. *J Environ Eng* 136(2): 227 – 237. doi: 10.1061/ASCEEE.1943-7870.0000147.

Wilkins D, Leung MHY, Lee PKH. (2017). Microbiota fingerprints lose individually identifying features over time. *Microbiome* 5: 1. doi: 10.1186/s40168-016-0209-7.

Williams MM, Domingo JWS, Meckes MC. (2005). Population diversity in model potable water biofilms receiving chlorine or chloramine residual. *Biofouling* 21(5 – 6): 279 – 288. doi: 10.1080/08927010500452695.

Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. (2012). Human gut microbiome viewed across age and geography. *Nature* 486(7402): 222 – 227. doi: 10.1038/nature11053.

Yu J, Kim D, Lee T. (2010). Microbial diversity in biofilms on water distribution pipes of different materials. *Water Sci Technol* 61(1): 163 – 171. doi: 10.2166/wst.2010.813.

# CHAPTER 2. IMPACT OF WATER CHEMISTRY, PIPE MATERIAL AND STAGNATION ON THE BUILDING PLUMBING MICROBIOME

Pan Ji, Jeffrey Parks, Marc Edwards, Amy Pruden

## ABSTRACT

A unique microbiome establishes in the portion of the potable water distribution system within homes and other buildings (i.e., building plumbing). To examine its composition and the factors that shape it, standardized cold water plumbing rigs were deployed at the treatment plant and in the distribution system of five water utilities across the U.S. Three pipe materials (copper with lead solder, CPVC with brass fittings or copper/lead combined pipe) were compared, with 8 hour flush cycles of 5 minutes to simulate typical daily use patterns. High throughput Illumina sequencing of 16S rRNA gene amplicons was employed to profile and compare the resident bulk water bacteria and archaea. The utility, location of the pipe rig, pipe material and stagnation all had a significant influence on the plumbing microbiome composition, but the utility source water and treatment practices were dominant factors. Examination of 21 water chemistry parameters suggested that the total chlorine concentration, pH, P, $SO4^{2-}$ and Mg were associated with the most of the variation in bulk water microbiome composition. Disinfectant type exerted a notably low-magnitude impact on microbiome composition. At two utilities using the same source water, slight differences in treatment approaches were associated with differences in rare taxa in samples. For genera containing opportunistic pathogens, Utility C samples (highest pH of 9-10) had the highest frequency of detection for *Legionella* spp. and lowest relative abundance of *Mycobacterium* spp. Data were examined across utilities to identify a true universal core, special core, and peripheral organisms to deepen insight into the physical and chemical factors that shape the building plumbing microbiome.

**KEYWORDS**

Microbiome, drinking water, building plumbing, water chemistry, opportunistic pathogens, *Legionella*, *Mycobacteria*

## INTRODUCTION

Drinking water systems are far from sterile environments, and recent application of molecular methods has revealed surprising diversity in composition (Berry et al., 2006; Holinger et al., 2014) and function (Gomez-Alvarez et al., 2012). The microbial ecology of drinking water systems is now understood to play a critical role in a wide range of economic, water management, and health problems, including microbial-induced corrosion (Burleigh et al., 2014; Wagner and Chamberlain, 1997), nitrification in chloraminated systems (Regan et al., 2003; Wolfe et al., 1990), and waterborne disease (Leclerc et al., 2002; Reynolds et al., 2008).

The portion of the drinking water distribution system within homes and other buildings (i.e., building plumbing or premise plumbing) creates a unique niche for microbial proliferation (Li et al., 2010; Prévost et al., 2014). Relative to the main water distribution system, the surface area to volume ratio is high (Prévost et al., 1997), the water is warm and stagnant during much of the day (National Research Council, 2006), and disinfectant can decay more rapidly (Nguyen et al., 2012), contributing to more prevalent regrowth of microorganisms relative to water mains. These distinctions are important because building plumbing represents the final gateway for exposure of consumers to the microbes inhabiting their drinking water, such as opportunistic pathogens (OPs) (e.g., *Legionella pneumophila*, *Mycobacterium avium*, *Pseudomonas aeruginosa*) (Wang et al., 2012a). Nonetheless, drinking water regulations are generally focused on water leaving the treatment plant and in the main distribution system and not at the point-of-use where exposure actually occurs.

OPs are now recognized as the primary source of waterborne disease outbreak in developed countries (Brunkard et al., 2011; Hilborn et al., 2013; Pruden et al., 2013). In contrast to traditional fecal pathogens, OPs are native to freshwater and drinking water systems (Falkinham III et al., 2015) and can proliferate in the building plumbing environment (Ashbolt, 2015; Feazel et al., 2009; Felfö Idi et al., 2010; van der Wielen et al., 2013; van Ingen et al., 2009). Also, many OPs are relatively tolerant of disinfectants (Le Dantec et al., 2002; Sidari et al., 2014), and broader strategies such as managing the microbial ecology, i.e. a “probiotic” approach, will ultimately be needed to effectively control OPs, as opposed to reliance on disinfectants alone

(Wang et al., 2013). Thus, a firm understanding of the microbial ecology of building plumbing systems is needed, and recent advances in the application of next generation sequencing technology (Lozupone et al., 2008) provides the opportunity to advance towards this goal. Next generation sequencing has successfully captured driving influence of on-site monochloramine to bacterial ecology in a hospital's hot water system (Baron et al., 2014; Baron et al., 2015) and also served to map point-of-use tap water microbiome across seventeen cities along the Arkansas and lower Mississippi rivers (Holinger et al., 2014).

While these pioneering studies provide a glimpse of the vast diversity of microbial communities inhabiting drinking water systems, there are many practical challenges to build a mechanistic understanding of the factors that shape the building plumbing microbiome. First of all, buildings are highly complex, with each essentially representing a unique array of pipe materials (Nguyen et al., 2012), configurations (World Health Organization, 2006), and flow-patterns (National Research Council 2006). A question of particular importance is to what degree the building plumbing microbiome is shaped by the influent water quality versus the building plumbing environment itself. Or, to what extent building owners have control over the building plumbing microbiome relative to water utilities. These are very difficult questions to answer based on field studies, as the inherent complexity of building plumbing precludes isolation of factors hypothesized to govern microbial community composition.

This study addresses the above challenges through installation of standardized building plumbing rigs at five water utilities, located in the eastern U.S., and characterization via Illumina sequencing of 16S rRNA gene amplicons. In order to capture the effects of changes in water quality during passage through water mains, one rig was located at or near the drinking water treatment plant and a second rig was installed in a distal portion of the distribution system. The rigs themselves facilitated comparison of pipe material (copper with lead solder, CPVC with brass fitting, copper/lead combined pipe), with uniform flow cycles (five minutes flow every eight hours) to simulate typical stagnation in household daily water usage (i.e. sleeping hours at night, working hours during daytime). The composition of the building plumbing microbiome was compared across rigs to gain insight into the relative influence of the source water chemistry and treatments, pipe material, water age, and stagnation events.

## MATERIALS AND METHODS

**Building plumbing rig design and sample collection.** Two standardized rigs were installed at each of five potable water utilities (A, B, C, D and E) in the eastern portion of the U.S. At each utility, one rig directly received water from the drinking water treatment plant (WTP) and the other was located in the distribution system (DS) at water ages of ≈ 4.5, 6.5, 0.75, 0.6 and 3 days, respectively (Parks et al., 2014). Each rig was constructed with three pipe materials in triplicate: copper with lead solder (referred to as "Copper"), plastic with brass fitting (referred to as "CPVC"), copper/ lead combined pipe (referred to as "Copper/lead"). All rigs were operated under identical flow conditions, mimicking typical household water usage patterns (i.e., 10 mL/min for 10 minutes, followed by 7 hours 50 minutes stagnation). The rigs had been acclimated for about 1 year at the time of sampling.

Utility personnel assisted in sampling the rig water between Nov.27$^{th}$ 2012 and Jan.16$^{th}$ 2013. A detailed sampling protocol and pre-assembled sampling kit was provided to each utility to ensure uniformity of sampling. Sample bottles were autoclaved 1000 mL polypropylene wide mouth square bottles (Nalgene$^{TM}$, U.S. Plastic Corp., Lima, OH) containing 26.7µL of 0.5g/mL sodium thiosulfate and 1.4 mL 0.5M EDTA (pH=8.0) added as preservatives (both sterilized using 0.22 µm filter) to limit damage of microbes and their DNA by any disinfectants or metals in the water.

For each sampling event (3 pipe materials in triplicate and 1 influent for each rig location, a total of 10 samples/rig-event), about 700 mL first flush bulk water samples were collected on three consecutive mornings or flow cycles (Batches 1, 2, 3). Sampling was conducted just prior to scheduled flushing for maximal impact of stagnation. One field blank was included for each sampling batch by transferring 700 mL autoclaved nanopure water, included with the kit, into a sterile sampling bottle after taking 10 samples. Once collected, samples were shipped on ice overnight to Virginia Tech, where they were immediately stored at 4°C and processed within 24 hours upon receipt.

**Site location and water chemistry characterization.** All five utilities received surface water as source water. Distinguishing features about the utilities, such as water treatment approach and geographic information, are reported in Table A.1 and Table A.10. In particular, Utilities A-D used chlorine to provide secondary residual in the distribution system, while Utility E used chloramines. Utilities A and B shared the same source water and both applied conventional treatment processes, with the distribution system at Utility A more than double in size than that of Utility B (Parks et al., 2014), which provided the opportunity to examine the effect of post-source water factors (see Discussion). Six to seven basic water chemistry parameters were measured throughout the study and on site at the time of Batch 1 sample collection, including: temperature, pH, free chlorine, total chlorine, turbidity, ammonia (chloraminated system only), and corrosion current (copper/lead combined pipes). Prior to Batch 1 sampling, background water chemistry data, including total organic carbon (TOC), $SO_4^{2-}$, $NO_3^-$, concentration of metal ions were measured at Virginia Tech over a 3 month period (Parks et al., 2014).

**Sample processing and DNA extraction.** Water samples were filtered through 0.22 μm-pore-size sterile mixed cellulose ester filters (GSWP047S0, EMD Millipore, Billerica, MA). The filters were fragmented using flame-sterilized tweezers and placed in 2 mL Lysing Matrix A tubes (MP Biomedicals, Solon, OH). A filter blank was taken at the end of each filtration round by loading it onto the manifold and turning the vacuum pump on for the average time required for samples. Tubes were stored at -80 °C, and later subject to DNA extraction using the FastDNA® SPIN Kit (MP Biomedicals) according to manufacturer instruction. A tube blank, which received only reagents and no sample, was included for each DNA extraction round.

**Illumina sequencing.** All samples were subjected to PCR amplification (Caporaso et al., 2010) using universal bacterial/archaeal primer set 515f/806r, which targets V4 region of 16S rRNA gene (Caporaso et al., 2012). Sample preparation followed the Earth Microbiome Project 16S rRNA Amplification Protocol (Earth Microbiome Project, 2015). Minor changes included using Molecular Biology Grade Water (Quality Biological, Gaithersburg, MD), and QIAquick PCR Purification Kit (QIAGEN, Valencia, CA). Pooled samples were submitted to the Virginia Bioinformatics Institute for paired-end 250 cycle Illumina sequencing (VBI GRL MiSeq sequencing service, Blacksburg, VA). Eight out of 300 samples were precluded due to low yield

in PCR products. Field blank, filter blank and tube blank samples were pooled on a "maximum volume" basis instead of equal molar criteria (maximum volume of other samples in the same lane). During the preparation process, all samples experienced identical freeze-thaw cycles.

**Data Analysis.** All original DNA sequences and metadata have been deposited to QIITA, under Study ID 10251. PANDAseq was used for joining paired-end sequence reads (Masella et al., 2012) in Oracle VM Virtual Box version 4.2.12. Stitched reads were pre-filtered based on length (252-255bp) before processed with the Quantitative Insights Into Microbial Ecology (QIIME) version 1.8.0 (Caporaso et al., 2010), following an open source online protocol (http://qiime.org/tutorials/tutorial.html, accessed 17 July 2015). The *pick_de_novo_otus.py* script was used to conduct: 1) *de novo* operational taxonomic unit (OTU) picking, using method uclust_ref (Edgar, 2010) with cutoff value of 3%, 2) taxonomy assignment of generated OTUs against Greengenes 13_5 reference database (McDonald et al., 2012), 3) phylogenetic tree construction by FastTree 2.1.3 (Price et al., 2010), and 4) OTU table construction. Singletons (defined as an OTU represented by 1 sequence, and appears only once in the whole OTU table) were removed from the OTU table prior to downstream analysis. A total of 35.7 million sequences were retrieved from 292 water samples, with a minimum of 29, 238 sequences per sample. Rarefaction to 29,238 sequences was applied to all samples before downstream analysis to minimize impact of uneven sequencing depth.

The Shapiro-Wilk normality test was applied for each water chemistry parameter. Kruskal-Wallis analysis was chosen to assess variance across different utilities. Both tests were performed in R version 3.0.2 (R Core Team, 2013). Principal Component Analysis (PCA) was applied for comparing water chemistry parameters across different samples (Primer 6, version 6.1.13). Jackknifed beta diversity based on both unweighted and weighted UniFrac distance matrices (Lozupone and Knight, 2005) were calculated in QIIME to compare microbial composition of different samples. The unweighted UniFrac distance matrix considers presence/absence of each operational taxonomic unit (OTU), while the weighted version includes relative abundance information. Emperor (Vazquez-Baeza et al., 2013) was used to visualize jackknifed beta diversity distance matrices. Analysis of similarity (ANOSIM) and similarity percentage (SIMPER) (both in Primer 6, version 6.1.13) were used to compare

similarity/dissimilarity of sample microbiome from the same utility/ utility pair. Indicator species were determined for each utility at species level ({indicspecies}, R). Adonis (permutational multivariate analysis of variance using distance matrices, {vegan}, R) was applied to explore potential impact of single factor or water chemistry parameter. BEST (Bio-Env+Stepwise) analysis (Primer 6, version 6.1.13) was applied to identify a "best" possible combination of water chemistry parameters that explained the largest portion of variation in microbial community composition across samples (Clarke, 1993). This "BEST" set was further chosen for Canonical correspondence analysis (CCA, {vegan} package, R) with taxonomy Table A.t species level. Statistical significance was set at $p<0.05$.

## RESULTS

**Factors Influencing Water Chemistry.** PCA provided a comprehensive comparison of water chemistry data across the utilities (Fig. 2.1) and revealed general trends, indicating that: 1) local water chemistry was distinct at each utility; 2) water chemistry of DS rig samples was distinct from that of WTP rig samples; 3) water chemistry changed during stagnation; and 4) water chemistry changed with pipe materials. The "utility" is an aggregate factor that includes properties of source water, water treatment and distribution process.

ANOSIM results indicated that utility, rig location, and pipe material all contributed significantly to the overall variance in water chemistry data across samples (alpha = 0.05, Table A.3). Among these three factors, utility resulted in the highest global R statistic of 0.713, compared to 0.332 for location of rig at WTP or in DS (nested under utility), and 0.448 for pipe material (nested under utility.rig). Comparing each influent to corresponding effluents, 8-hour stagnation did not have a measureable impact on water chemistry, regardless of pipe material ($P = 0.18$, nested under utility. rig). The impact of stagnation was noted to be non-uniform for each pipe material (i.e., concentrations of Pb, Cu, Zn), which could have masked the overall impact of stagnation on water chemistry.

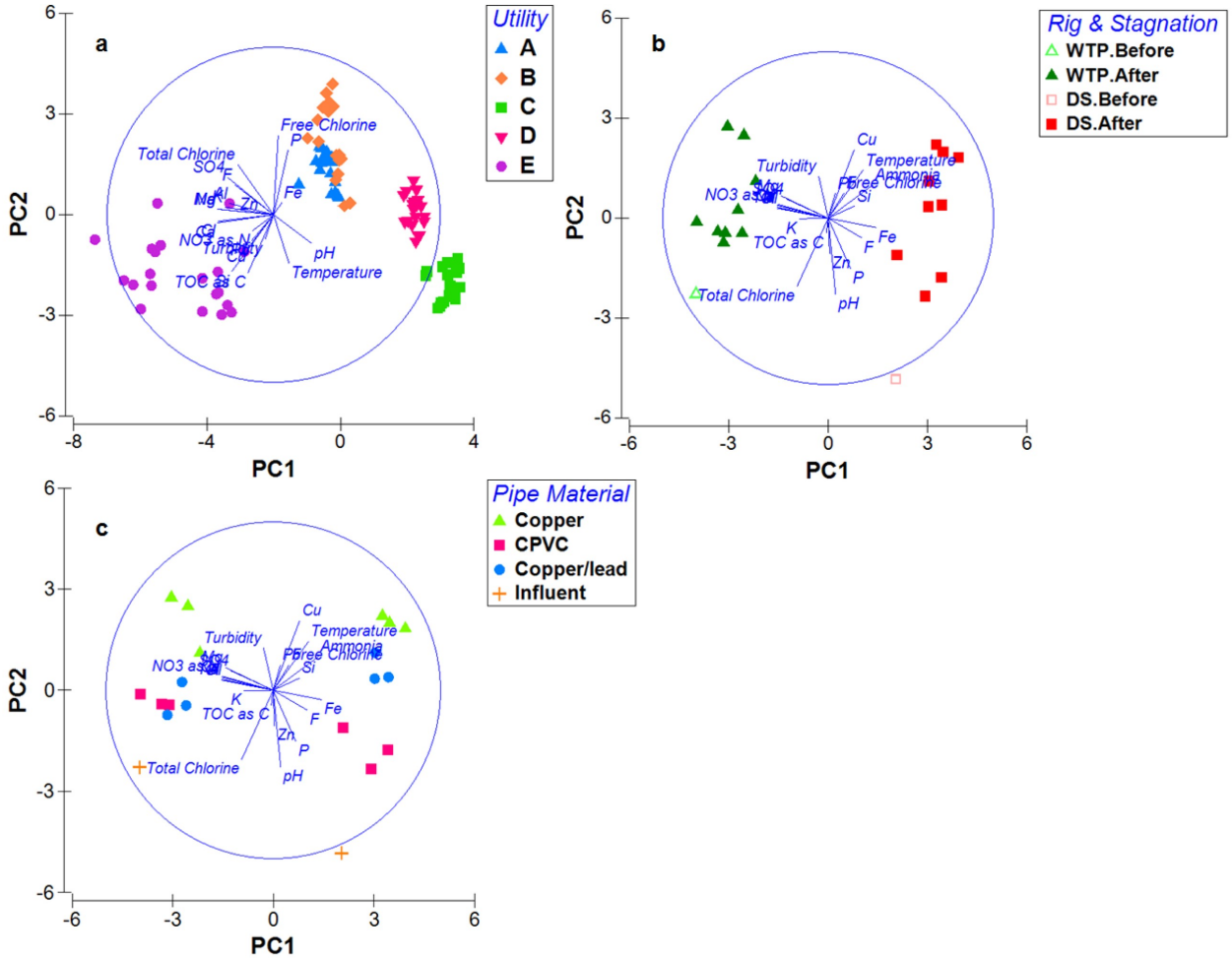


**Fig. 2.1.** Dissimilarity in water chemistry of samples from different utilities, rig locations, and pipe materials.

PCA plot of water chemistry data. Each point represents a water sample, with proximity of points in 2-D space indicative of relative similarities. Principle component (PC) 1 and PC2 are combinations of water chemistry variables that best explained variation among samples. a. samples from five utilities (n=100 samples), color and shape coded based on utility location; PC1 and PC2 explained 44.2% and 27.3% variation, respectively. b. samples from Utility E (n=20), shape coded by rig location and color coded by stagnation stage; PC1 and PC2 explained 46.5% and 16.3% each. c. samples from Utility E, color and shape coded by pipe material; PC1 and PC2 explained 46.5% and 16.3% each.

All water chemistry parameters varied across utilities ($P < 0.05$, Table A.4), except for Zn. Total chlorine concentration was highest at Utility B (1.16±0.38 mg/L) and lowest at Utility C (0.23±0.28 mg/L). In most cases, disinfectant decay was apparent in the influents to the DS rigs

relative to the WTP rigs based on examination of data 3-month prior to sampling (Fig. A.1). An exception was the higher total chlorine levels observed in the influent to the DS rig relative to WTP rig at Utility B (~ 0.5 mg/L), which is likely the result of fluctuation in water treatment and chlorine dosing, along with the fact that there was a 6.5 day water age delay between the DS and WTP rigs. The highest pH occurred at Utility C (9-10), with other utilities in the more typical, EPA recommended range of 6.5-8.5 (Fig. A.2). Certain parameters, including concentrations of Pb, Cu, and Zn after stagnation, were mostly influenced by pipe material, rather than utility (Table A.4). For instance, Pb concentration was usually highest in copper/lead combined pipe (992.6±1773.9 ppb, with 190.0±498.9 ppb in copper pipe and 3.1±3.7 ppb in CPVC pipe); Cu concentration was highest in copper pipe (252.6±268.5 ppb, with 139.2±163.6 ppb in copper/lead combined pipe and 33.8±34.7 ppb in CPVC pipe); and Zn concentration was highest in CPVC pipe (141.3±126.8 ppb, with 32.9±79.0 ppb for copper pipe and 19.2±32.1 ppb for copper/lead combined pipe). "Spikes" of turbidity mostly (9 out of 10 values larger than 1 NTU) occurred in copper pipe and copper/lead combined pipe.

Utility A and B, which shared the same source water, had the most similar water chemistry relative to the other utilities (Fig. 2.1), with a few distinctions noted. For instance, Utility B had higher total chlorine concentration (1.16±0.38 mg/L), but lower pH (7.3±0.1) compared to Utility A (total chlorine 0.73±0.24 mg/L, pH 8.6±0.1). Also, Utility B had about twice the phosphorus concentration as Utility A (250.9±21.2 ppb vs 135.1±5.7 ppb).

**Microbiome Composition.** *Microbial community composition*. A total of 3 archaeal and 37 bacterial phyla were detected across all samples, with 0.002% archaeal sequences, 99.3% bacterial sequences and 0.7% unclassified sequences. Four dominant (i.e., >1% relative abundance) bacterial phyla accounted for 96.9% of the total sequences: *Actinobacteria* (17.6%), *Bacteroidetes* (2.6%), *Cyanobacteria* (8.2%) and *Proteobacteria* (68.6%) (Fig. 2.2).

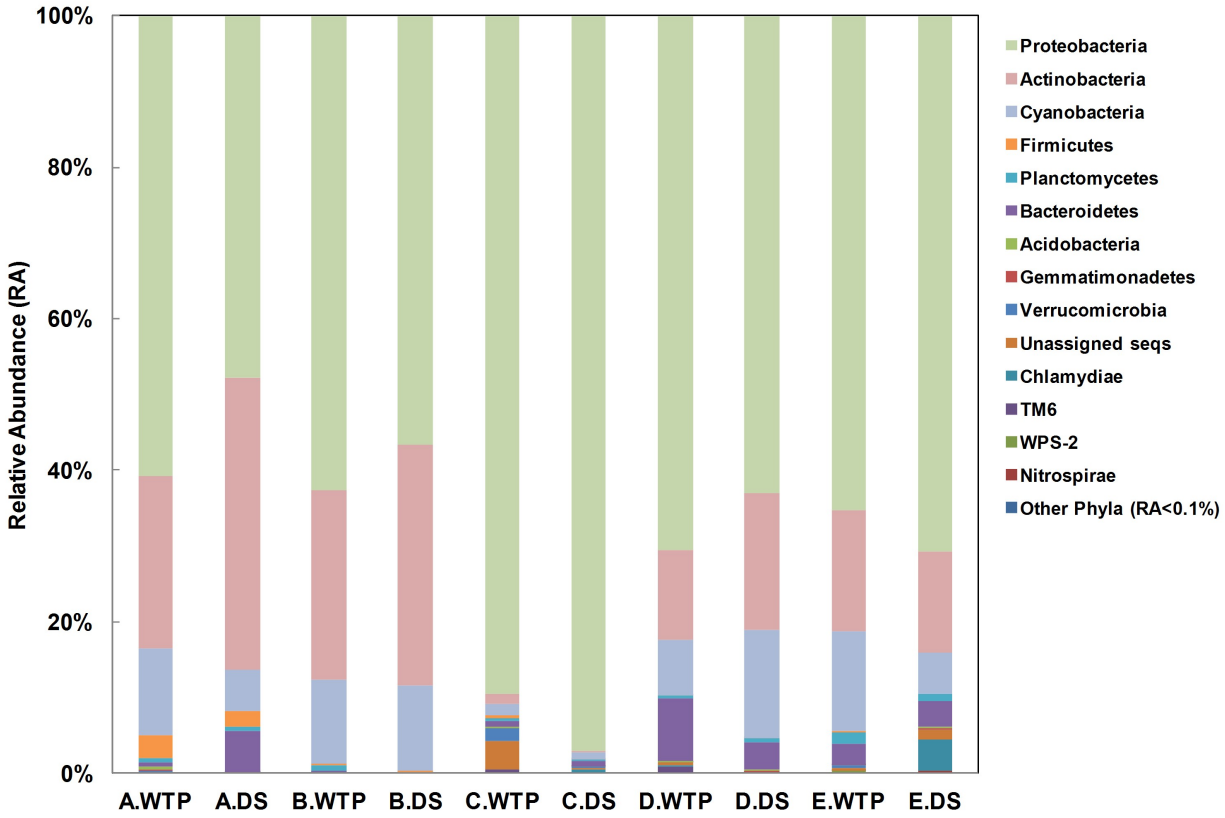


**Fig. 2.2. Microbiome taxonomy composition of samples from each rig (phylum level).**

Data were combined across all 27 pipe samples for each rig. Relative abundance was calculated as the ratio of sequences. Phyla with relative abundance less than 0.1% were combined into "Other Phyla (RA < 0.1%)".

*Genera Containing Opportunistic Pathogens (OPs).* Genera containing OPs were widely detected in samples across the five utilities (Table 2.1). *Legionella* spp. were detected in samples from all five utilities, except from Utility D, with the highest frequency of detection in Utility C samples. At Utility C, *Legionella* spp. were detected in 29 out of 30 WTP rig samples and 18 out of 29 DS rig samples, whereas all detections in Utility E occurred in DS rig samples.

**Table 2.1. Frequency of detection of genera containing OPs across the five utilities** (n= 60, 54, 59, 60, and 59 samples for Utility A, B, C, D and E, respectively)**.**

| | A | B | C | D | E |
|---|---|---|---|---|---|
| ***Legionella* spp.** | 1.7% | 7.4% | 78.0% | 0.0% | 13.6% |
| ***Mycobacterium* spp.** | 100% | 100% | 98.3% | 100% | 100% |
| ***Pseudomonas* spp.** | 96.7% | 48.1% | 86.4% | 70.0% | 93.2% |

*Mycobacterium* spp. were detected in all samples except one. The relative abundance of *Mycobacterium* spp. was noted to be 2 orders of magnitude lower at Utility C ($7.4\times10^{-4}$) relative to the other four utilities (>0.14). The frequency of detection of *Pseudomonas* spp. was > 70% of samples across all utilities, with the exception of Utility B. Utility B samples had the lowest relative abundance (0.01%) of *Pseudomonas* spp., while highest relative abundance was observed in Utility A samples (1.54%).

*Comparison of the Microbial Community Composition*. 3-D beta diversity plots constructed from both unweighted and weighted UniFrac distance matrices illustrated several key conclusions about microbial community composition (Fig. 2.3): 1) samples from the same utility were more similar to each other in terms of presence/absence of specific OTUs, relative to other utilities, 2) within the same utility, samples from the WTP and DS rig were distinct, 3) within the same rig, pipe material and stagnation both had an influence. It is of interest to note that the distinction in microbial community composition across the utilities was most apparent based on the unweighted UniFrac matrix (which considers presence/absence of OTUs), not the weighted UniFrac distance matrix (which considers relative abundance in addition to presence/absence of each OTU). However, both distance matrices indicated similar 3-D patterns with respect to effect of rig location and pipe material within each utility/rig.

ANOSIM pair-wise tests (Table A.5) based on weighted and unweighted UniFrac distance matrices confirmed the above 3-D patterns. A global R value greater than 0.75 indicates microbial communities strongly different in composition, whereas an R value between 0.5 to 0.75 indicates some compositional overlap. For unweighted UniFrac distance matrix comparisons, a global R value greater than 0.75 was obtained for all paired utility comparisons, except for B vs D. For weighted UniFrac distance matrix comparisons, paired comparisons between C and the other four had an R>0.75, while all utility pairs among A, B, D, and E had an R<0.5, suggesting large compositional overlap.

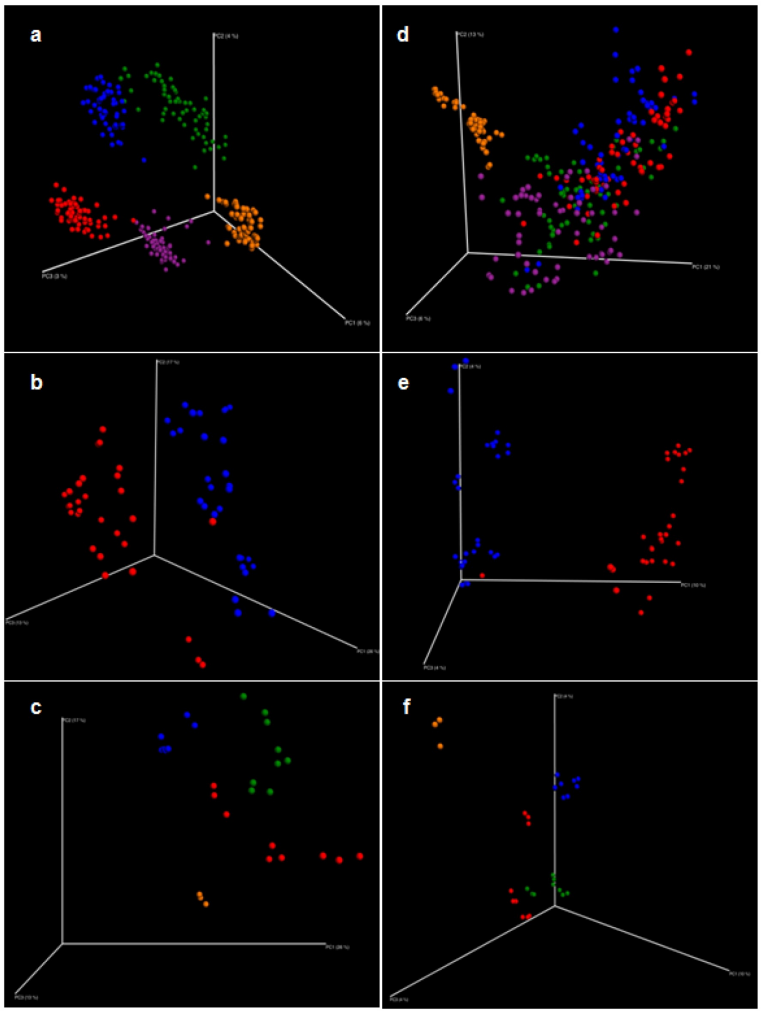


**Fig. 2.3. Dissimilarity in microbiome composition of samples from different utilities, rig locations, and pipe materials.**

3-D beta diversity plots derived from jackknifed unweighted (a, b, c) and weighted (d, e, f) UniFrac distance matrices, color coded by: 1) utility (a and d), all samples (n = 60, 54, 59, 60, 59, color = red, blue, yellow, green, purple for A, B, C, D, E, respectively); 2) rig location (b and e), Utility E samples (WTP in blue, n=29; DS in red, n = 30); 3) pipe material and stagnation (c and f), Utility E, WTP rig samples (n = 9, 9, 9, 3, color = blue, red, green, brown for Copper, CPVC, Copper/lead, and influent, respectively).

Utility, rig location, pipe material and stagnation each had a significant impact on microbiome composition, with utility as the master factor (Table 2.2). Utility, rig, pipe material and stagnation together explained the greatest dissimilarity in microbiome across samples, followed by utility and rig, then utility only. Within each set of utility samples, relative importance of rig

location and pipe material varied (Table A.6). All factors showed larger magnitude of impact when calculated based on weighted UniFrac distance matrix (Table 2.2). Sample Batch had a statistically significant, but low magnitude impact on microbiome composition, and thus Batch 1 samples were further examined to explore the association between microbiome composition and water chemistry.

**Table 2.2. Impact of various factors on the microbiome across all samples (n=292).**

| Factor | Strata | Unweighted UniFrac | | Weighted UniFrac | |
|---|---|---|---|---|---|
| | | $R^2$ | P | $R^2$ | P |
| **Utility & Rig & Pipe & Stagnation** | | 0.339 | 0.001 | 0.702 | 0.001 |
| **Utility.rig** | | 0.220 | 0.001 | 0.520 | 0.001 |
| **Utility** | | 0.157 | 0.001 | 0.387 | 0.001 |
| **Rig** | Utility | 0.013 | 0.001 | 0.034 | 0.001 |
| **Pipe Material** | Utility.rig | 0.014 | 0.001 | 0.030 | 0.001 |
| **Stagnation** | Utility.rig | 0.005 | 0.001 | 0.015 | 0.001 |
| **Batch** | Utility.rig.pipe | 0.005 | 0.001 | 0.003 | 0.008 |
| **Disinfectant type** | | 0.032 | 0.001 | 0.066 | 0.001 |

Adonis analysis was applied using package "vegan" from R, with permutation = 999. Unweighted UniFrac considers presence/absence of each OTU, while weighted UniFrac also considers relative abundance of each OTU. "Strata" were defined based on sampling design as the overarching factor, with the subsequent hierarchy of factors derived according to relative magnitudes of impact on the microbiome. Further permutations were constrained within a given stratum. For example, when Utility.rig (utility and rig location) was set as the stratum for the purpose of examining the impact of stagnation, samples from the same Utility.rig were pooled and randomized, but samples across different Utility.rig combinations were not.

*Microbial classes driving similarity and dissimilarity.* SIMPER (Similarity Percentages- species contribution, Primer 6) analysis was applied to explore microbial classes that drive microbial similarity within the same utility, and dissimilarity across different utilities (Table A.7). In general, *Alphaproteobacteria*, *Actinobacteria*, *4C0d-2*, *Betaproteobacteria*, and *Gammaproteobacteria* were the main contributors in driving microbiome similarity within each utility and dissimilarity across different paired utilities. These were also top 5 abundant classes among all sequences combined.

*Candidate indicator taxa for each utility.* Indicator taxa are genera (or lowest taxonomy levels) whose presence/absence and relative abundance respond to particular environmental variables or specific environments and therefore best characterize a particular group of samples. With respect to utility location, indicator taxa were: 1) mostly found in that particular utility and 2) present in the majority of samples from that utility (Dufrene and Legendre, 1997). For each utility, the top 10 candidate indicator taxa with the highest indicator value and statistical significance <0.05 are reported in Table A.8. For Utility B, all indicator genera fell within the top 52% of genera ranked from highest to lowest relative abundance. For the other four utilities, all but one of the indicator genera were among the top 20% of genera ranked according to relative abundance.

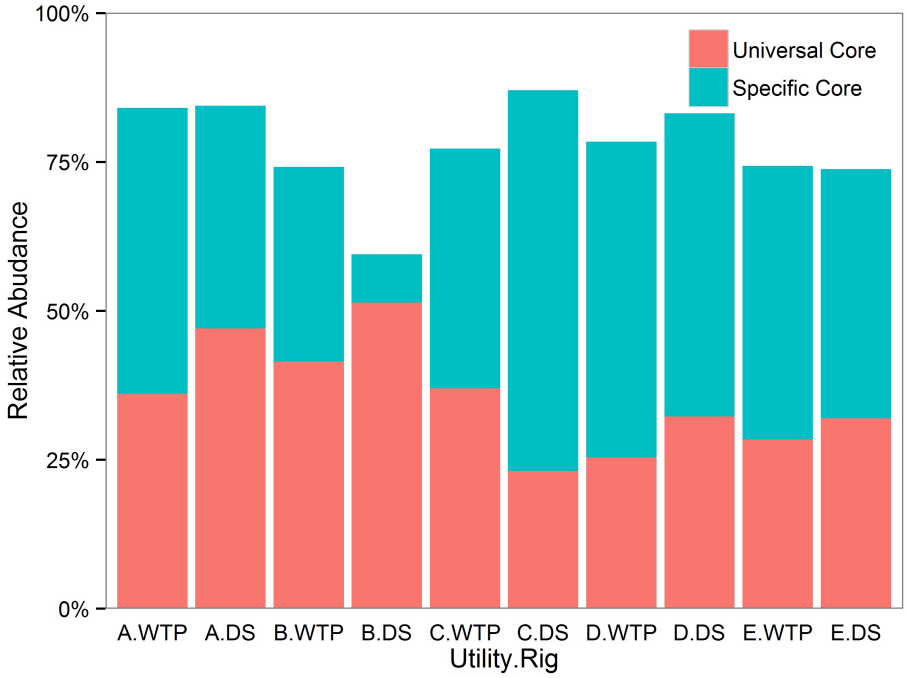


**Fig. 2.4. Core OTU comparison across each rig location at each utility.**

Relative abundance was calculated by normalizing number of core OTU sequences to the total number of sequences within specific Utility.Rig combination. The universal core is defined as OTUs shared among all samples, while the specific core consists of OTUs shared within each Utility.Rig, but not across all samples.

*"Core" microbiome.* One question of interest is if a "core" drinking water microbiome can be identified (Proctor and Hammes, 2015). The core microbiome is often defined as OTUs shared above a selected threshold percentage of samples within a certain category (*e.g.* samples from

same utility, samples collected on same day) (QIIME, core microbiome, http://qiime.org/scripts/compute_core_microbiome.html, accessed 17 July 2015). In this study, the threshold percentage was set at 100% and category set defined as utility and rig location together (termed "Utility.Rig"). Core microbiome for each utility & rig combination were further classified into two parts- a universal core (shared among all samples from five utilities combined), and a specific core (OTUs shared among samples within a specific Utility.Rig category, but not in all samples). Size of core microbiome varied, on average 78.0±14.1% of the total sequences (Fig. 2.4).

A second question is whether the "core" microbiome for each utility/rig category are essentially representative of the microbiome of all samples. The Mantel test was applied to compare UniFrac distance matrices constructed from core microbiome by Utility.Rig and the total microbiome. There was a significant and strong positive correlation between corresponding distance matrices (unweighted UniFrac, R statistic = 0.837; weighted UniFrac, R statistic = 0.907; P = 0.001), implying potential of a core microbiome capturing main patterns across all samples.

**Table 2.3. Association between microbiome and lumped water chemistry variables.**

| | **Unweighted UniFrac** | **Weighted UniFrac** |
|---|---|---|
| **"BEST" combination** | pH, Mg | Total chlorine, pH, P, $SO_4^{2-}$, Mg |
| **Rho statistic** | 0.741 | 0.501 |

BEST analysis was conducted using Primer 6. Rho statistic represents "association strength", ranging from 0 to 1 with rho>0.5 to be generally considered a strong association. Permutation = 99.

**Association between microbiome and water chemistry.** *Association between microbiome and individual water chemistry parameters*. Most water chemistry parameters were associated with significant differences in microbiome composition (both weighted and unweighted UniFrac distance matrices), although the strength of association was much less than the aggregate association of the microbiome with utility (Table A.9). Among 21 parameters, turbidity did not have a statistically significant association with the microbiome composition, while $NO_3^-$ and Pb only indicated a significant association when using the unweighted UniFrac distance matrix

(presence/absence of microbes). The remaining parameters all were more strongly associated with differences in the microbiome composition when applied to the weighted UniFrac distance matrix. Only six parameters had a strength of association larger than 0.1 (weighted UniFrac).

*Association between microbiome and multiple parameters*. The influent and effluent waters of the pipe rigs, including rigs at Utilities A and B, which shared the same original source water, were distinct in both water chemistry and microbiome composition. Of all possible combinations of water chemistry parameters, total chlorine concentration, pH, phosphorus (P), sulfate ($SO_4^{2-}$) and magnesium (Mg) together explained the greatest dissimilarity of microbiome composition of corresponding samples across the five utilities (weighted UniFrac distance matrix) (Table 2.3). These parameters also displayed higher individual association with the microbiome (Table A.9). If only considering presence/absence of microbes (unweighted UniFrac distance matrix), pH and Mg played predominant roles (Table 2.3).

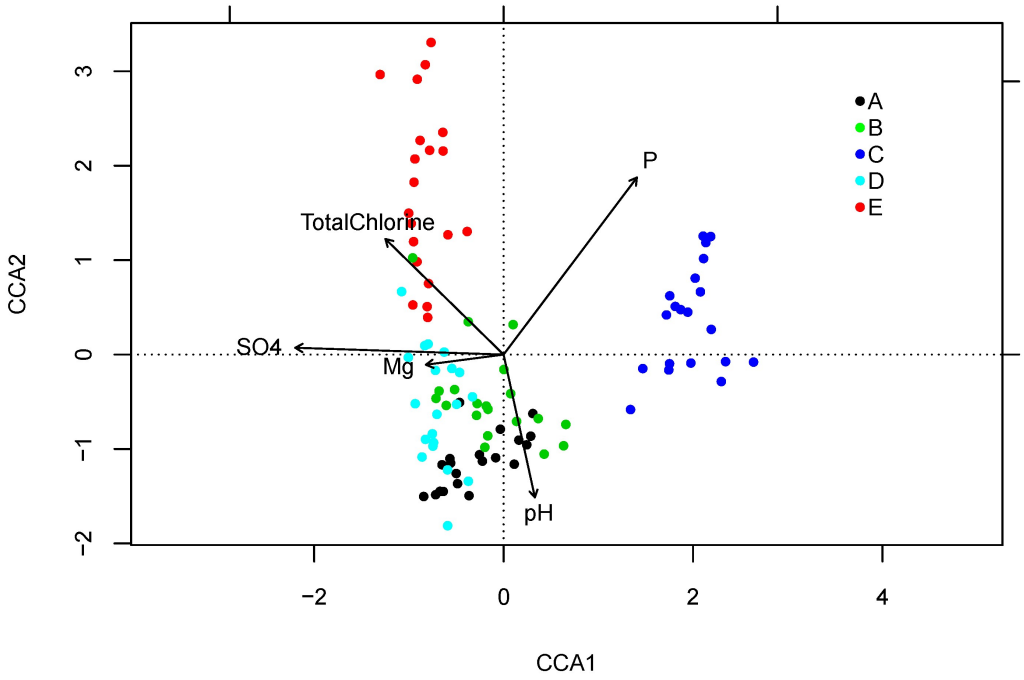


**Fig. 2.5. Microbiome composition (genus level) in association with "BEST" water chemistry parameters of Batch 1 samples.**

Each point represents microbiome of one sample. CCA1 and CCA2 each explained 43.4% and 21.1% of all five constrained axes generated by Canonical Correspondence Analysis (CCA).

Canonical Correspondence Analysis (CCA) was applied to visualize how microbial taxonomy composition (at genus level) was associated with the "BEST" set of water chemistry parameters, including total chlorine, pH, P, $SO_4^{2-}$, and Mg (Fig. 2.5). The portion of variance in microbiome taxonomy that could be explained by the "BEST" set was 20.2%. The pattern in Fig. 2.5 was in accordance with water chemistry data based on the fact that: 1) Utility E samples were differentiated from others due to highest total chlorine concentration; 2) Utility C samples had the lowest P (near to 0 compared to others at hundreds ppb level), highest pH (up to 10 compared to others in neutral range), lowest Mg (less than 10% compared to others) and lowest $SO_4^{2-}$ (~20% compared to Utilities A, B and E); and 3) Utility D samples had ~16% of the $SO_4^{2-}$ concentration compared to Utilities A and B, with samples from these three utilities varying along a gradient of $SO_4^{2-}$.

## DISCUSSION

**Utility is the overarching factor for both water chemistry and microbiome composition.** The unique, standardized and replicated pipe rig design employed in this study enabled a controlled comparison of several variables on the ultimate microbial composition at the tap. Among the four main factors that were the focus of this study: utility, rig location, pipe material and stagnation all had a significant influence on water chemistry and microbial composition. Among these, the utility itself had an overarching influence (Table 2.2 and Table A.3). Notably, though Utilities A and B shared the same source water, the chemistries and microbiome compositions of the corresponding rigs were distinct, emphasizing the importance of water treatment and distribution in shaping the building plumbing microbiome. This finding is consistent with a recent lab-scale study, in which water generated from simulated distribution systems representing different pipe materials, disinfectants, and water ages was the dominant influence on the composition of the microbial community and occurrence of OPs in simulated household water heaters, rather than the warm, stagnant environment of the heater itself (Wang et al., 2015).

A recent Netherlands survey on distribution systems corresponding to four different groundwater treatment plants (flushed water sample, Illumina sequencing of 16S rRNA gene amplicons) also

indicated that utility, over distance (0.4-35 km to treatment plant) and time (3 sampling events in 17 months), was the predominant factor driving microbial community composition in drinking water (Roeselers et al., 2015). Interestingly, potable water is not subject to secondary chlorination in the Netherlands, indicating that other subtleties of the aquifer and water chemistry, besides disinfectants, can drive significant differences in the corresponding microbial community.

In addition to source water properties and the water treatment/distribution processes, the term "utility" includes local geology, regional watershed characteristics, climate, and weather or other events at the time of sampling. Geographical location is likely to also be important (Fierer, 2008; Martiny et al., 2006), in addition to historical events. One study proposed possible historical pig manure application and agricultural runoff to explain the concurrence of high nitrate concentration and significant enrichment in *Erysipelothrix* spp. (Holinger et al., 2014).

**Impact of source water vs water treatment and distribution process.** Source water type (surface vs groundwater) and ratios of blends of different source waters was found to influence abundance and proportion of certain taxa, *e.g.* sulfur metabolism related microbes were enriched in groundwater sources (Holinger et al., 2014; Pinto et al., 2014). In this study, all source waters were surface water (either river, lake, or reservoir); however, the water chemistry into and out of the building plumbing rigs still displayed distinct features, implying that the influence of water chemistry goes beyond broad classification of surface water versus groundwater.

Of special interest is whether the drinking water treatment and distribution process could override source-water characteristics in terms of the microbiome that shows up at the tap. Microbial composition has been observed to shift during the various stages of treatment from source water (Lin et al., 2014). One other study did indicate that moving downstream of the treatment application, the source water (surface water from River Murray) outweighed impact of different treatment processes (conventional coagulation, magnetic ion exchange resin plus conventional coagulation, magnetic ion exchange resin plus conventional coagulation and granular activated carbon, and membrane filtration) on microbial composition of the biofilm at a distance as short as 1km (Shaw et al., 2014).

Utilities A and B shared the same source water and similar water chemistry (Fig. 2.1) and their microbiome compositions (weighted UniFrac, Fig. 2.3) were correspondingly most similar amongst all utility pairs, indicating an influential role of source water. Nonetheless, there were no differences in the microbiome composition and water chemistry features between Utilities A and B, which likely related to differences in water treatment and distribution. Specifically: a) Utility A mostly used alum as coagulant with occasional application of ferric sulfate, while Utility B used ferric sulfate exclusively; b) for corrosion control, Utility A employed lime softening and caustic, while Utility B used blended phosphate and caustic; c) Utility A applied granular activated contactors along with traditional treatment, and UV disinfection as an extra safe-guard. The above differences most likely induced variation in several key water chemistry parameters (e.g. pH, total chlorine, phosphorus) for treated waters from Utilities A and B. For instance, elevated phosphorus level in Utility B water samples most likely arose from application of phosphate for corrosion control. The influenced parameters were all among the "BEST" set that identified parameters most strongly associated with differences in microbiome composition across utilities. While the microbiome compositions of Utilities A and B were similar by weighted UniFrac method (with relative abundance), they were distinct by unweighted UniFrac method (presence/absence only). This discrepancy indicates that the above differences in water treatment and distribution between the two utilities were especially influential on rare taxa, the importance of which should not be overlooked (e.g., OPs are typically "rare"). Thus, while source water is likely an overarching factor, water treatment and distribution management choices available to utilities alter the key water chemistry components and ultimate microbiome composition in the building plumbing.

**Influence of location in the distribution system, stagnation, pipe material**. Comparison of the WTP and DS rig samples was of particular interest, as this provides insight into the role of water age in shaping the building plumbing microbiome. In this study the water age entering the DS rigs ranged from 0.5-7 days. In a laboratory simulation of a water distribution system, water age (ranging from 1 to 10.2 days) was observed to significantly shape both the microbiome and key water chemistry parameters (disinfectant residual, total organic carbon and dissolved oxygen) (Masters et al., 2015; Wang et al., 2014a). This study confirmed that water quality at

the tap varied based on location in the distribution system, even with as little of difference of 0.5 d water age, and provided insight into the relative degrees of associated microbial shifts.

To the knowledge of the authors', this is the first controlled study of the impact of a typical building plumbing stagnation period, ~8 hr, on the microbial composition of the tap water. In terms of water chemistry, stagnation dramatically influenced free chlorine, and selectively influenced Pb, Cu and Zn concentrations for certain pipe materials. Water samples after stagnation yielded 6-13 more phyla compared to corresponding influents (before/after stagnation: A 10/22, B 9/16, C 13/26, D 14/20, E 16/29). The increase suggested potential "seeding" from building plumbing biofilm, or regrowth of rare species above the detection limit, likely as a result of disinfectant decay and the magnified influence of biofilms in the small diameter pipe. Overnight stagnation was reported to result in cell regrowth and a shift in 50-100% of the microbial community composition in drinking water samples (bulk water) from different households in a non-chlorinated water system (Lautenschlager et al., 2010). In this study, a significant effect of stagnation was observed in all utilities, except for Utility B. It was noted that samples after stagnation were still more closely related to the corresponding influent to the rig than to the stagnated water from pipes of the same type from other rigs. Similar spatial stability in drinking water treatment plant and distribution systems have been reported in other studies (Roeselers et al., 2015; Eichler et al., 2006; Hwang et al., 2012; Pinto et al., 2012).

Pipe material exhibited a significant effect on 4 out of 21 water chemistry parameters: Pb, Cu, Zn and turbidity (Table A.4). These parameters are highly pipe material-specific: as Pb was released mostly in copper/lead combined pipe, less in copper pipe (i.e. lead solder) and only slightly in CPVC pipe; Cu was released mostly in copper/lead combined pipe and copper pipe; and Zn was released mainly from brass fitting on CPVC pipe. Turbidity likely was a byproduct associated with metal particles and biofilm released from metal pipes. Pipe material appeared to be strongest determinant of the microbiome composition at Utility E, outweighing the impact of the rig location at the WTP or in the DS (Table A.6). Pipe materials are known to influence biofilms in terms of biofilm density (Niquette and Servais, 2000), bacterial diversity (Jang et al., 2011), formation potential (Yu et al., 2010) and formation rate (Lehtola et al., 2004), which could indirectly influence bulk water microbiome composition.

**What are the predominant water chemistry parameters?** Among the "BEST" water chemistry parameter set across multiple utilities, total chlorine concentration (Hwang et al., 2012), pH (Pinto et al., 2012), P and $SO_4^{2-}$ together (Pinto et al., 2014) have all been previously reported as strongest determinant(s) of the microbial composition of drinking water. A few studies examine the importance of Mg to the drinking water microbiome, and we speculate that Mg in this study is an indicator of salinity (Fig. A.3), which is widely reported to be the strongest determinant of microbial community structure in aquatic systems (Herlemann et al., 2011; Lozupone and Knight, 2007; Wu et al., 2006). CCA analysis further suggested the potential for complex interrelationships among water chemistry parameters (Fig. 2.5), indicating "cooperative", rather than individual, influence on the microbiome. For example, disinfectant efficacy varies among target microbes, while pH governs the relative proportions of hypochlorous acid versus hypochlorite, which also have differing efficacies.

Surprisingly, TOC was not among the "BEST" variables identified as having influence on the microbial composition. The assimilable portion of the TOC (LeChevallier et al., 1991; Miettinen et al., 1997) and phosphorus (Miettinen et al., 1997; Lehtola et al., 2001) have been noted as important limiting nutrients for microbial growth in drinking water distribution systems. In a recent study of simulated household water heaters, total heterotrophic plate count bacteria correlated with TOC, but ranged within the same order of magnitude (2,000 μg/L to 15,000 μg/L) (Williams et al., 2015). In the present study, TOC was generally low (1035±603 μg/L) across all utilities and thus likely was not in an ideal range to capture variability.

In contrast to previous reports of the dominant effect of chloramine versus chlorine in shaping microbial composition (Gomez-Alvarez et al., 2012; Baron et al., 2015; Hwang et al., 2012; Wang et al., 2014b), we observed limited effect of disinfectant type on microbiome composition of the building plumbing microbiome relative to the other factors investigated (Table 2.2). Greater dissimilarity in microbiome composition occurred among chlorinated systems, relative to paired comparisons with the chloraminated utility (Table A.5). Disinfectant type appeared to impose stronger influence on rare taxa (Holinger et al., 2014), an effect that would have limited resolution with the weighted UniFrac distance matrix approach (relative abundance included in

calculation). Again, while disinfectant type clearly has an effect, this study indicates that the utility as a whole is the overarching factor. We further speculate that the precise predominant water chemistry parameters vary within each utility, and are not necessarily the same across utilities. It is also noteworthy that the most influential water chemistry variables might change through seasons or due to operational changes (Pinto et al., 2014).

**Several abundant taxa comprised the majority of microbiome.** A total of 22, 16, 26, 20, and 29 different phyla were recovered for Utility A, B, C, D, and E, respectively, while previous drinking water studies have recovered: 9 to 16 (Pinto et al., 2012), 26 (Holinger et al., 2014), 26 (Wang et al., 2014b), 11 (Zeng et al., 2013), and 15 (Liu et al., 2014). Consistent with prior drinking water surveys, the majority of total sequences were assigned to a few dominant phyla (Holinger et al., 2014; Pinto et al., 2014). The dominant phyla observed in this study, *Actinobacteria*, *Bacteroidetes*, *Cyanobacteria* and *Proteobacteria*, have been commonly reported in other drinking water distribution systems (Holinger et al., 2014; Eichler et al., 2006; Poitelon et al., 2009). Such high similarity in dominant phyla across multiple study sites could be attributed to unique aspects of the drinking water environment, including similar micro-environments (i.e. standardized rig setting, water treatment plant) (Holinger et al., 2014). Within the Proteobacteria phylum, other studies have also noted the overall predominance of the *Alphaproteobacteria* class in water mains (Gomez-Alvarez et al., 2012; Douterelo et al., 2013; Douterelo et al., 2014), and a shift within proteobacteria to *Alphaproteobacteria* during winter season (Pinto et al., 2014).

**Genera containing opportunistic pathogens (OPs).** Amplicon sequences obtained from Illumina sequencing are short (~250 bp, without primers or barcodes), thus they have limited taxonomic resolution and cannot definitively identify pathogens. However, several genera known to contain OPs were identified, and examining these patterns could provide insight into the behavior of pathogenic members. Water chemistry, microbiome composition, and occurrence of genera containing OPs appeared to have a complex interrelationship (Falkinham III et al., 2015; Wang et al., 2013). *Legionella* is acid-tolerant (Nguyen et al., 1991) and most isolates have been obtained from environmental sources in the pH range of 2.7-8.3 (Dimitriadi et al., 2014). Further, it has been reported in prior studies that *Legionella pneumophila* has reduced

viability and cultivability at higher pH (Ohno et al., 2003; States et al., 1987). Thus, hypothetically, the high pH level at Utility C (9-10) would have negatively influenced the occurrence of *Legionella* spp. In contrast, Utility C had the highest frequency of detection of *Legionella* spp. One possible explanation is that the relatively low disinfectant residual levels, which have been associated with high frequency of detection of *Legionella* spp. (Edagawa et al., 2008; Wang et al., 2012b), offset the potential inhibitory effects of elevated pH. *Legionella* spp. had higher frequency of detection in WTP rig samples (28 out of 29) relative to DS rig samples (18 out of 30), while disinfectant residual level in the DS rig was ~0.5 mg/L lower than that of WTP rig samples. This suggested there are not necessarily simple linear relationships between *Legionella* spp. and isolated water chemistry parameters in real-world pipes. It is interesting that Utility C samples also had the lowest relative abundance of *Mycobacterium* spp., indicating the potential for competition between the two OPs-containing genera, or environmental selection favoring *Legionella* spp. over *Mycobacterium* spp., as higher pH is expected to result in more effective inactivation of the latter (Le Dantec et al., 2002).

**Nitrogen-cycling bacteria.** Ammonia-oxidizing bacteria (AOB) and nitrite-oxidizing bacteria are a nuisance in chloraminated drinking water distribution systems because they catalyze decay of chloramine disinfectant (Regan et al., 2003; Wolfe et al., 1990; Regan et al., 2002). In this study, nitrite-oxidizing bacteria were primarily *Nitrospira* spp., which was also an indicator genus for the chloraminated utility (Utility E). Predominance of *Nitrospira* spp. over *Nitrobacter* spp. (latter in 5 out of 59 samples) might be explained by advantages of k-strategists over r-strategists in oligotrophic water: the former exhibits a low maximum specific growth rate, but is adapted to low nitrite and oxygen concentrations (Nogueira et al., 2006). Common AOB genera, including *Nitrosomonas*, *Nitrosoccocus* and *Nitrosospira* (Li et al., 2007), were not detected. However, the potential AOB-related family, *Nitrosomonadaceae*, was detected at a relative abundance of (1.3±0.8%) across all Utility E samples, which was the highest frequency of detection among the five utilities. *Candidatus Brocadia*, mainly detected in freshwater/estuary (Sun et al., 2014), was detected in 1 sample only, suggesting the possibility of autotrophic bacteria using nitrite to oxidize ammonia under anoxic conditions (Tal et al., 2006; van Niftrik and Jetten, 2012).

**Candidate indicator taxa**. Indicator genera suggested potentially important metabolic processes occurring in the rigs. *Hydrogenophaga*, an indicator genus at Utility B, is a hydrogen-oxidizer (Williams et al., 1989). At Utility C, *Limnobactor* is associated with sulfur-oxidation, with certain species able to oxidize thiosulfate (Lu et al., 2011). *Methylocaldum* spp. are mesophilic to thermophilic obligate methanotrophs, using methane and methanol as carbon and energy source (Dworkin et al., 2006). *Hyphomicrobium* spp. are methylotrophic microbes widely detected in various habitats including groundwater and fresh water (Holm et al., 1996).

Indicator taxa could also reflect source water background. As an indicator for Utility D, *Polaromonas* genus is psychrophilic and predominant in glacial, periglacial and high-elevation environments (Darcy et al., 2011), which is concordant with the source water feature and geographic location of Utility D. Also, indicator taxa provide preliminary insight into potential risk. In Utility C, *Legionella* spp., which contains *Legionella pneumophila*, was one of the indicator genera. At Utility A, *Staphylococcus* was the top ranked indicator genus. While *Staphylococcus* has high occurrence on human skin (Kloos and Schleifer, 1975), its absence from any of three field blanks implies little likelihood of originating from contamination by sampling personnel. A prior survey noted that 6% of individual rural western drinking water wells cultured positive for natural sources of *Staphylococcus aureus* (LeChevallier and Seidler, 1980).

**Quality control indicators for sampling low biomass water**. To maximize yield of DNA and consider repeatability of sampling procedures, each rig was sampled three times within a three-day period. Sampling batch exerted little to no influence when the microbiome composition was compared within each utility (Table A.6) and across utilities (Table 2.2). Long-term temporal stability is also sometime observed, as was the case in a 17-month study of an unchlorinated distribution systems in the Netherlands (Roeselers et al., 2015). However, a study in Michigan indicated dramatic seasonal patterns (Pinto et al., 2014).

Of all 39 blank samples (30 field blanks, 4 filter blanks and 5 tube blanks), only 3 field blank samples (all from WTP rig at Utility A) yielded detectable sequences: 943, 1162, and 1212 sequences (Table A.2). The proportion of blank sequences to the minimum number of sequences

obtained among samples was less than 7%. Thus, extent of potential "contamination" sequences was considered to be small in this study.

**Core microbiome insight into source of the building plumbing microbiome.** A question of interest is whether a healthy drinking water environment would "select" for a core microbiome. This study encompasses 10 different geographical locations (Utility.Rig), with the "universal core" defined as OTUs shared in all samples across the five utilities. The universal core contained five OTUs, three of which were also detected in all three field blanks that yielded sequences, but at very low abundance as mentioned above. A "special core" was defined as OTUs shared within samples from the same Utility.Rig combination, but not represented in the universal core. OTUs not belonging to either core are termed here as "peripheral". Size of the universal core and special core within each Utility.Rig category varied, together constituting over 55% of the total microbiome. Moreover, Mantel test results indicated that the core microbiome was representative of the total microbiome.

Existence of a true universal core, special core, and peripheral organisms implies potential differences in dispersal limitations. For instance, the universal core might be linked with high dispersal and colonization rate, high tolerance to stress, and the ability to compete with indigenous microbes (Martiny et al., 2006). Dispersal limitations were recently suggested for the planktonic portion of the water distribution system microbiome (Pinto et al., 2014). The composition of the drinking water microbiome can vary across different phases (e.g., bulk water, biofilm, loose deposits and suspended solids) (Liu et al., 2014), stages (e.g., source water, treatment plant, distribution system, built environment/building plumbing) (Proctor and Hammes, 2015), and seasons (Pinto et al., 2014). Such high complexity would inevitably lead to a core microbiome that shifts as a function of time and location (Pinto et al., 2014). The core microbiome concept could be useful in further studies in elucidating the driving factors that dictate the composition of the microbiome at the tap (Proctor and Hammes, 2015).

## CONCLUSION

This study employs a unique standardized building plumbing rig design to provide controlled, replicated comparison of potential factors contributing to the microbial composition in drinking water at the tap, questions that are confounded by the complexities faced in field studies. Overall, it was observed that the utility itself was the overarching factor in shaping the building plumbing microbiome, while location of the rig in the distribution system, plumbing material, and 8-hr stagnation events also had a significant influence. Total chlorine concentration, pH, P, $SO_4^{2-}$ and Mg together explained greatest variation in microbiome at multi-utility level, however within each utility such "BEST" sets might change. This study suggests that factors under control of the utility, including physical/chemical properties of the water and prior treatments, drive the composition of the building plumbing microbiome, including the occurrence of opportunistic pathogens. However, factors under control of building owners also clearly have an influence. These findings have important implications for water engineering and management, helping lay the groundwork needed to identify critical factors that may be manipulated in the future to beneficially manage the building plumbing microbiome.

## ACKNOWLEDGEMENTS

The views presented here do not represent those of the sponsors. We thank all five water utilities and related personnel on their willingness to participate and generous help on sampling. We also thank Brian Badgley for lab supplies, Lisa Belden, Matthew Becker and Jenifer Walke for assistance on data analysis.

## AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: AP MAE PJ. Performed the experiments: PJ JP. Analyzed the data: PJ. Wrote the paper: PJ AP MAE.

## REFERENCES

Ashbolt N. (2015). Microbial contamination of drinking water and human health from community water systems. *Curr Envir Health Rpt* 2(1): 95 – 106.

Baron JL, Vikram A, Duda S, Stout JE, Bibby K. (2014). Shift in the microbial ecology of a hospital hot water system following the introduction of an on-site monochloramine disinfection system. *PLoS ONE* 9(7): e102679. doi: 10.1371/journal.pone.0102679.

Baron JL, Harris JK, Holinger EP, Duda S, Stevens MJ, Robertson CE, *et al*. (2015). Effect of monochloramine treatment on the microbial ecology of *Legionella* and associated bacterial populations in a hospital hot water system. *Syst Appl Microbiol* 38(3): 198 – 205.

Berry D, Xi C, Raskin L. (2006). Microbial ecology of drinking water distribution systems. *Curr Opin Biotechnol* 17(3): 297 – 232.

Brunkard JM, Ailes E, Roberts VA, Hill V, Hilborn ED, Craun GF, et al. (2011). Surveillance for waterborne disease outbreaks associated with drinking water—United States, 2007-2008. *Morb Mortal Wkly Rep* 60 (ss12): 38 – 68.

Burleigh TD, Gierke CG, Fredj N, Boston PJ. (2014). Copper tube pitting in Santa Fe municipal water caused by microbial induced corrosion. *Materials* 7(6): 4321 – 4334.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7(5): 335 – 336.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6: 1621-1624.

Clarke KR. (1993). Non-parametric multivariate analyses of changes in community structure. *Aust J Ecol* 18: 117 – 143.

Darcy JL, Lynch RC, King AJ, Robeson MS, Schmidt SK. (2011). Global distribution of *Polaromonas* phylotypes – evidence for a highly successful dispersal capacity. *PLoS ONE* 6(8): e23742.

Dimitriadi D, Velonakis E. (2014). Detection of *Legionella* spp. From domestic water in the prefecture of Arta, Greece. *J Pathog* 407385. doi: 10.1155/2014/407385.

Douterelo I, Sharpe RL, Boxall JB. (2013). Influence of hydraulic regimes on bacterial community structure and composition in an experimental drinking water distribution system. *Water Res* 47(2): 503 – 516.

Douterelo I, Husband S, Boxall JB. (2014). The bacteriological composition of biomass recovered by flushing an operational drinking water distribution system. *Water Res* 54(1): 100 – 114.

Dufrene M, Legendre P. (1997). Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol Monogr* 67(3): 345 – 366.

Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stachebrandt E. (Eds.) (2006). The Prokaryotes (third edition): a handbook on the biology of bacteria, Volume 5. New York: Springer Science & Business Media. pp 275-278.

Earth Microbiome Project: 16S rRNA amplification protocol version 4_13. Available at: ftp://ftp.metagenomics.anl.gov/data/misc/EMP/SupplementaryFile1_barcoded_primers_515F_806R.txt. Accessed 17 July 2015.

Edagawa A, Kimura A, Doi H, Tanaka H, Tomioka K, Sakabe K, et al. (2008). Detection of culturable and nonculturable *Legionella* species from hot water systems of public buildings in Japan. *J Appl Microbiol* 105(6): 2104 – 2114.

Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19): 2460 – 2461.

Eichler S, Christen R, Höltje C, Westphal P, Bötel J, Brettar I, et al. (2006). Composition and dynamics of bacterial communities of a drinking water supply system as assessed by RNA- and DNA-based 16S rRNA gene fingerprinting. *Appl Environ Microbiol* 72(3): 1858 – 1872.

Falkinham III JO, Hilborn ED, Arduino MJ, Pruden A, Edwards MA. (2015). Epidemiology and ecology of opportunistic premise plumbing pathogens: *Legionella pneumophila*, *Mycobacterium avium*, and *Pseudomonas aeruginosa*. *Environ Health Perspect* 123(8): 749 – 758. doi: 10.1289/ehp.1408692.

Feazel LM, Baumgartner LK, Peterson KL, Frank DN, Harris JK, Pace NR. (2009). Opportunistic pathogens enriched in showerhead biofilms. *PNAS* 106(38): 16393 – 16399.

Felfö Idi T, Heéger Z, Vargha M, Márialigeti K. (2010). Detection of potentially pathogenic bacteria in the drinking water distribution system of a hospital in Hungary. *Clin Microbiol Infec* 16(1): 89 – 92.

Fierer N. (2008). Microbial biogeography: patterns in microbial diversity across space and time. In: K. Zengler (editor). Accessing uncultivated microorganisms: from the environment to organisms and genomes and back. Washington DC: ASM Press; pp 95-115.

Gomez-Alvarez V, Revetta RP, Santo Domingo JW. (2012). Metagenomic analysis of drinking water receiving different disinfection treatments. *Appl Environ Microbiol* 78(17): 6095 – 6102.

Herlemann DPR, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF. (2011). Transitions in bacterial communities along the 2000 km salinity gradient the Baltic sea. *ISME J* 5: 1571 – 1579.

Hilborn ED, Wade TJ, Hicks L, Garrison L, Carpenter J, Adam E, et al. (2013). Surveillance for waterborne disease outbreaks associated with drinking water and other nonrecreational water—United States, 2009 – 2010. *Morb Mortal Wkly Rep* 62(35): 714 – 720.

Holinger EP, Ross KA, Robertson CE, Stevens MJ, Kirk Harris J, Pace NR. (2014). Molecular analysis of point-of-use municipal drinking water microbiology. *Water Res* 49(1): 225 – 235.

Holm NC, Gliesche CG, Hirsch P. (1996). Diversity and structure of *Hyphomicrobium* populations in a sewage treatment plant and its adjacent receiving lake. *Appl Environ Microbiol* 62(2): 522 – 528.

Hwang C, Ling F, Anderson GL, LeChevallier MW, Liu W-T. (2012). Microbial community dynamics of an urban drinking water distribution system subjected to phases of chloramination and chlorination treatments. *Appl Environ Microbiol* 78(22): 7856 – 7865.

Jang H-J, Choi Y-J, Ka J-O. (2011). Effects of diverse water pipe materials on bacterial communities and water quality in the annular reactor. *J Microbiol Biotechnol* 21(2): 115 – 123.

Kloos WE, Schleifer KH. (1975). Isolation and characterization of Staphylococci from human skin (II), Descriptions of four new species: *Staphylococcus warneri, Staphylococcus capitis, Staphylococcus hominis, and Staphylococcus simulans*. *Int J Syst Bacteriol* 25(1): 62 – 79.

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

LeChevallier MW, Seidler RJ. (1980). *Staphylococcus aureus* in rural drinking water. *Appl Environ Microbiol* 39(4): 739 – 742.

LeChevallier MW, Schulz W, Lee RG. (1991). Bacterial nutrients in drinking water. *Appl Environ Microbiol* 57: 867 – 862.

Leclerc H, Schwartzbrod L, Dei-Cas E. (2002). Microbial agents associated with waterborne disease. *Crit Rev Microbiol* 28(4): 371 – 409.

Le Dantec C, Duguet J-P, Montiel A, Dumoutier N, Dubrou S, Vincent V. (2002). Chlorine disinfection of atypical mycobacteria isolated from a water distribution system. *Appl Environ Microbiol* 68(3): 1025 – 1032. doi:10.1128/AEM.68.3.1025-1032.2002.

Lehtola MJ, Miettinen IT, Vartiainen T, Myllykangas T, Martikainen PJ. (2001). Microbially available organic carbon, phosphorus, and microbial growth in ozonated drinking water. *Water Res* 35(7): 1635 – 1640.

Lehtola MJ, Miettinen IT, Keinänen MM, Kekki T, Laine O, Hirvonen A, et al. (2004). Microbiology, chemistry and biofilm development in a pilot drinking water distribution system with copper and plastic pipes. *Water Res* 38(17): 3769 – 3779.

Li B, Irvin S, Baker K. (2007). The variation of nitrifying bacterial population sizes in a sequencing batch reactor (SBR) treating low/mid/high concentrated wastewater. *J Environ Eng Sci* 6: 651 – 663.

Li D, Li Z, Yu J, Cao N, Liu R, 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.

Lin W, Yu Z, Zhang H, Thompson IP. (2014). Diversity and dynamics of microbial communities at each step of treatment plant for potable water generation. *Water Res* 52(1): 218 – 230.

Liu G, Bakker GL, Li S, Vreeburg JH, Verberk JQ, Medema GJ, et al. (2014). Pyrosequencing reveals bacterial communities in unchlorinated drinking water distribution system: an integral study of bulk water, suspended solids, loose deposits, and pipe wall biofilm. *Environ Sci Technol* 48(10): 5467 – 5476.

Lozupone CA, Knight R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71(12): 8228 – 8235.

Lozupone CA, Knight R. (2007). Global patterns in bacterial diversity. *PNAS* 104: 11436 – 11440.

Lozupone CA, Knight R. (2008). Species divergence and the measurement of microbial diversity. *FEMS Microbiol Rev* 32(4): 557 – 578.

Lu H, Sato Y, Fujimura R, Nishizawa T, Kamijo T, Ohta H. (2011). *Limnobacter litoralis* sp. nov., a thiosulfate-oxidizing, heterotrophic bacterium isolated from a volcanic deposit, and emended description of the genus *Limnobacter*. *Int J Syst Evol Microbiol* 61: 404 – 407.

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, et al. (2006). Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* 4(2): 102 – 112.

Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. (2012). PANDAseq: PAired-eND Assembler for Illumina sequences. *BMC Bioinform* 13: 31.

Masters S, Wang H, Pruden A, Edwards M. (2015). Redox gradients in distribution systems influence water quality, corrosion, and microbial ecology. *Water Res* 68(1): 140 – 149.

McDonald D, Price MN, Goodrich JA, Nawrocki EP, DeSantis TZ, Probst A, et al. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 6: 610 – 618.

Miettinen IT, Vartiainen T, Martikainen PJ. (1997). Phosphorus and bacterial growth in drinking water. *Appl Environ Microbiol* 63(8): 3242 – 3245.

National Research Council. (2006). Drinking water distribution systems: assessing and reducing risks. Washington, D.C.: National Academies Press; pp: 317 – 320.

Niquette P, Servais P. (2000). Impacts of pipe materials on densities of fixed bacterial biomass in a drinking water distribution system. *Water Res* 34(6): 1952 – 1956.

Nguyen MH, Stout JE, Yu VL. (1991). Legionellosis. *Infect Dis Clin North Am* 5(3): 561 – 584.

Nguyen C, Elfland C, Edwards M. (2012). Impact of advanced water conservation features and new copper pipe on rapid chloramines decay and microbial regrowth. *Water Res* 46: 611 – 621.

Nogueira R, Melo LF. (2006). Competition between *Nitrospira* spp. and *Nitrobacter* spp. in nitrite-oxidizing bioreactors. *Biotechnol Bioeng* 95(1): 169 – 175.

Ohno A, Kato N, Yamada K, Yamaguchi K. (2003). Factors influencing survival of *Legionella pneumophila* serotype 1 in hot spring water and tap water. *Appl Environ Microbiol*. 2003; 69(5): 2540-2547.

Parks J, Edwards M. (2014). Non-intrusive methodology for assessing lead and copper corrosion. Water Research Foundation. Available at: http://www.waterrf.org/PublicReportLibrary/4317.pdf. Accessed 17 July 2015.

Pinto AJ, Xi C, Raskin L. (2012). Bacterial community structure in the drinking water microbiome is governed by filtration processes. *Environ Sci Technol* 46(16): 8851 – 8859.

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

Poitelon JB, Joyeux M, Welte B, Duguet JP, Prestel E, DuBow MS. (2009). Variations of bacterial 16S rDNA phylotypes prior to and after chlorination for drinking water production from two surface water treatment plants. *J Ind Microbiol Biotechnol* 37: 117 – 128.

Prévost M, Rompré A, Baribeau H, Coallier J, Lafrance P. (1997). Service lines: their effect on microbiological quality. *J Am Water Works Assoc* 89(7): 78 – 91.

Prévost M, Besner MC, Laurent P, Servais P. (2014). Emerging issues of biological stability in drinking water distribution systems. In: van der Kooij D, and van der Wielen PWJJ (Ed.). Microbial growth in drinking water supplies: problems, causes, control and research news. London, UK: IWA publishing; pp 265-271.

Price MN, Dehal PS, Arkin AP. (2010). FastTree 2- approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5(3): e9490. doi: 10.1371/journal.pone.0009490.

Proctor CR, Hammes F. (2015). Drinking water microbiology- from measurement to management. *Curr Opin Microbiol* 33: 87 – 94.

Pruden A, Edwards MA, Falkinham III JO. (2013). State of the science and research needs for opportunistic pathogens in premise plumbing. Water Research Foundation. Available: http://www.waterrf.org/PublicReportLibrary/4379.pdf.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013; Available: http://www.R-project.org/.

Regan JM, Harrington GW, Noguera DR. (2002). Ammonia- and nitrite-oxidizing bacterial communities in a pilot-scale chloraminated drinking water distribution system. Appl *Environ Microbiol* 68(1): 73 – 81.

Regan JM, Harrington GW, Baribeau H, De Leon R, Noguera NR. (2003). Diversity of nitrifying bacteria in full-scale chloraminated distribution systems. *Water Res* 37: 197 – 205.

Reynolds KA, Mena KD, Gerba CP. (2008). Risk of waterborne illness via drinking water in the United States. *Rev Environ Contam Toxicol* 192: 117 – 158.

Roeselers G, Coolen J, van der Wielen PWJJ, Jaspers MC, Atsma A, de Graaf B, et al. (2015). Microbial biogeography of drinking water: patterns in phylogenetic diversity across space and time. *Environ Microbiol* 17(7): 2505 – 2514. doi: 10.1111/1462-2920.12739.

Shaw JLA, Monis P, Fabris R, Ho L, Braun K, Drikas M, et al. (2014). Assessing the impact of water treatment on bacterial biofilms in drinking water distribution systems using high-throughput DNA sequencing. *Chemosphere* 117: 185 – 192.

Sidari III FP, Stout JE, Duda S, Grubb D, Neuner A. (2014). Maintaining *Legionella* control in building water systems. *J Am Water Works Assoc* 106(10): 24 – 32.

States SJ, Conley LF, Towner SG, Wolford RS, Stephenson TE, McNamara AM, et al. (1987). An alkaline approach to treating cooling towers for control of *Legionella pneumophila*. *Appl Environ Microbiol* 53: 1775 – 1779.

Sun W, Xia C, Xu M, Guo J, Wang A, Sun G. (2014). Diversity and distribution of planktonic anaerobic ammonium-oxidizing bacteria in the Dongjiang River, China. *Microbiol Res* 169(12): 897 – 906.

Tal Y, Watts JEM, Schreier HJ. (2006). Anaerobic ammonium-oxidizing (anammox) bacteria and associated activity in fixed-film biofilters of a marine recirculating aquaculture system. *Appl Environ Microbiol* 72(4): 2896 – 2904.

Wagner D, Chamberlain AHL. (1997). Microbiologically influenced copper corrosion in potable water with emphasis on practical relevance. *Biodegradion* 8(3): 177 – 187.

Wang H, Edwards MA, Falkinham III JO, Pruden A. (2012a). Molecular survey of occurrence of *Legionella spp.*, *Mycobacterium spp.*, *Pseudomonas aeruginosa* and amoeba hosts in two chloraminated drinking water distribution systems. *Appl Environ Microbiol* 78(17): 6285 – 6294. doi: 10.1128/AEM.01492-12.

Wang H, Masters S, Hong Y, Stallings J, Falkinham JO, Edwards MA, et al. (2012b). Effect of disinfectant, water age, and pipe material on occurrence and persistence of *Legionella*, *Mycobacteria*, *Pseudomonas aeruginosa*, and two amoebas. *Environ Sci Technol* 46(21): 11566 – 11574.

Wang H, Edwards MA, Falkinham III JO, Pruden A. (2013). Probiotic approach to pathogen control in premise plumbing systems? A review. *Environ Sci Technol* 47: 10117−10128.

Wang H, Masters S, Edwards MA, Falkinham III JO, Pruden A. (2014a). Effect of disinfectant, water age, and pipe materials on bacterial and eukaryotic community structure in drinking water biofilm. *Environ Sci Technol* 48(3): 1426 – 1435.

Wang H, Proctor CR, Edwards MA, Pryor M, Santo Domingo JW, Ryu H, et al. (2014b). Microbial community response to chlorine conversion in a chloraminated drinking water distribution system. *Environ Sci Technol* 48(18): 10624 – 10633.

Wang H, Masters S, Falkinham III JO, Edwards MA, Pruden A. (2015). Distribution system water quality affects responses of opportunistic pathogen gene markers in household water heaters. *Environ Sci Technol* 49(14): 8416 – 8424. doi: 10.1021/acs.est.5b01538.

Williams A, Busse J, Goor M, Pot B, Falsen E, Jantzen E, et al. (1989). *Hydrogenophaga*, a new genus of hydrogen-oxidizing bacteria that includes *Hydrogenophaga flava* comb. nov. (formerly *Pseudomonas flava*), *Hydrogenophaga palleronii* (formerly *Pseudomonas palleronii), Hydrogenophaga pseudoflava* (formerly *Pseudomonas pseudoflava* and "*Pseudomonas carboxydoflava*"), and *Hydrogenophaga taeniospiralis* (formerly *Pseudomonas taeniospiralis*). *Int J Syst Bacteriol* 39(3): 319 – 333.

Williams K, Pruden A, Falkinham III JO, Edwards MA. (2015). Relationship between organic carbon and opportunistic pathogens in simulated glass water heaters. *Pathogens* 4(2): 355 – 372.

Wolfe RL, Lieu NI, Izaguirre G, Means EG. (1990). Ammonia-oxidizing bacteria in a chloraminated distribution system: seasonal occurrence, distribution, and disinfection resistance. *Appl Environ Microbiol* 56(2): 451 – 462.

World Health Organization. (2006). Health aspects of plumbing. World Health Organization and World Plumbing Council. Available at: http://www.who.int/water_sanitation_health/publications/plumbinghealthasp.pdf?ua=1. Accessed 17 July 2015.

Wu QL, Zwart G, Schauer M, Kamst-vanAgterveld MP, Hahn MW. (2006). Bacterioplankton community composition along a saline gradient of sixteen high-mountain lakes located on the Tibetan plateau, China. *Appl Environ Microbiol* 72(8): 5478 – 5485.

van der Wielen PW, van der Kooij D. (2013). Nontuberculous mycobacteria, fungi, and opportunistic pathogens in unchlorinated drinking water in the Netherlands. *Appl Environ Microbiol* 79(3): 825 – 834.

van Ingen J, Blaak H, de Beer J, de Roda Husman AM, van Soolingen D. (2009). Rapidly growing nontuberculous mycobacteria cultured from home tap and shower water. *Appl Environ Microbiol* 76(17): 6017 – 6019.

van Niftrik L, Jetten MSM. (2012). Anaerobic ammonium-oxidizing bacteria: unique microorganisms with exceptional properties. *Microbiol Mol Biol Rev* 76(3): 585 – 596.

Vazquez-Baeza Y, Pirrung M, Gonzalez A, Knight R. (2013). Emperor: A tool for visualizing high-throughput microbial community data. *GigaScience* 2(1): 16.

Yu J, Kim D, Lee T. (2010). Microbial diversity in biofilms on water distribution pipes of different materials. *Water Sci Technol* 61(1): 163 – 171.

Zeng DN, Fan ZY, Chi L, Wang X, Qu WD, Quan ZX. (2013). Analysis of the bacterial communities associated with different drinking water treatment processes. *World J Microbiol Biotechnol* 29: 1573 – 1584.

## APPENDIX A – SUPPLEMENTARY INFORMATION FOR CHAPTER 2

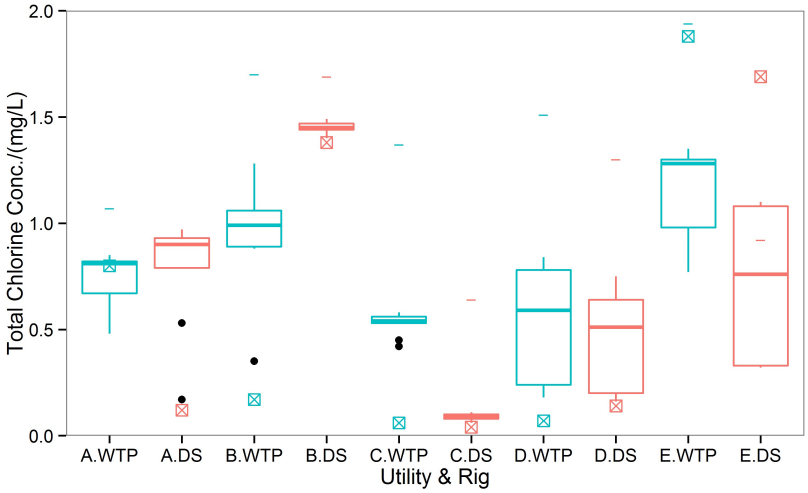


**Fig. A.1. Total chlorine concentrations of samples from water treatment plant (WTP) and distribution system (DS) rigs.** Each rig contains 9 pipe samples (triplicates of the three materials) following ~8hrs stagnation. Squares with cross inside are influent water samples, short lines represent average influent total chlorine concentration during 3-month prior to sampling event. Utilities A-D delivered chlorinated water, while Utility E delivered chloraminated water. The total chlorine concentration range was from 0.04-1.88 mg/L.

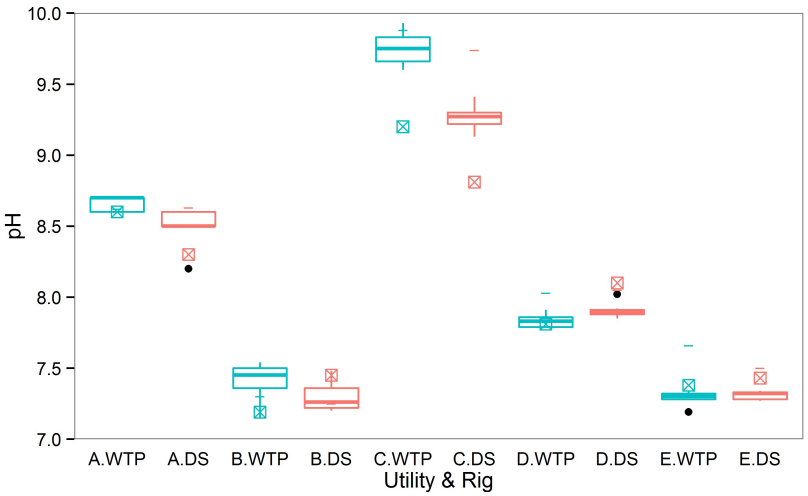


**Fig. A.2. pH of samples from water treatment plant (WTP) and distribution system (DS) rigs.** Each box contains 9 pipe samples from same rig (triplicates of the three materials) following ~8hrs stagnation. Squares with cross inside are influent water samples, while short lines are average influent pH values during 3-month prior to sampling.

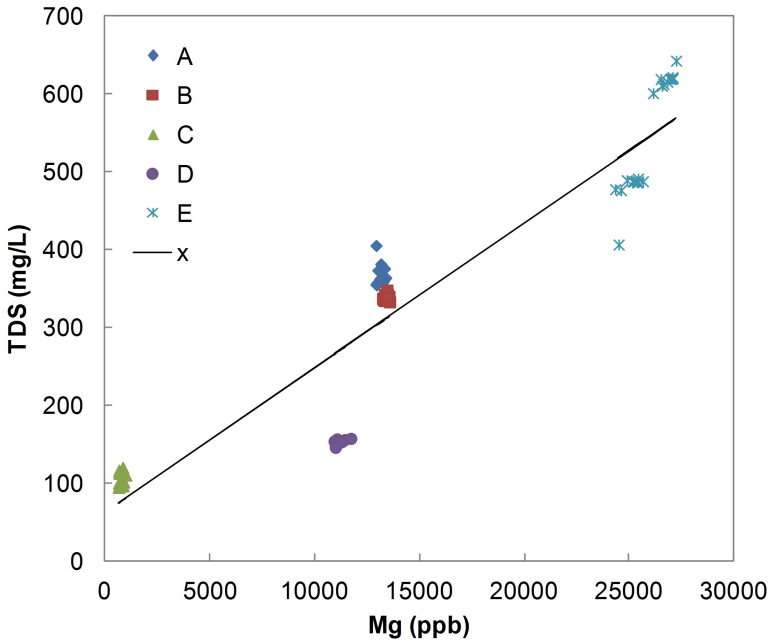


**Fig. A.3. Positive correlation between total dissolved solids and Mg concentrations.** Total dissolved solids were calculated using water chemistry data, representing salinity of drinking water.

**Table A.1. Geographical distances among five utilities.**

| | A | B | C | D | E |
|---|---|---|---|---|---|
| **A** | 0 | - | - | - | - |
| **B** | 12 | 0 | - | - | - |
| **C** | 844 | 823 | 0 | - | - |
| **D** | 373 | 380 | 1043 | 0 | - |
| **E** | 112 | 119 | 891 | 261 | 0 |

**Table A.2. Taxonomy composition of 3 field blanks from Utility A, WTP rig (genus level).** Sample names are Batch number when field blank was taken. Sequences number for each blank is in bracket. Minimum sequences per sample was 29, 238. Underlined are genera occurred in all three blanks.

| Taxon | 1 (943) | 2 (1162) | 3 (1212) |
|---|---|---|---|
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Ralstonia | 0.023 | 0.007 | 0.445 |
| k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas | 0.015 | 0.076 | 0.002 |
| k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__ | 0.120 | 0.002 | 0.004 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__ | 0.006 | 0.067 | 0.002 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Sphingomonas | 0.007 | 0.005 | 0.011 |
| k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Mycobacteriaceae;g__Mycobacterium | 0.004 | 0.010 | 0.004 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Erythrobacteraceae;g__ | 0.003 | 0.003 | 0.001 |
| k__Bacteria;p__Cyanobacteria;c__4C0d-2;o__MLE1-12;f__;g__ | 0.010 | 0.001 | 0.001 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__ | 0.003 | 0.003 | 0.001 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Sphingopyxis | 0.003 | 0.001 | 0.002 |
| Unassigned;Other;Other;Other;Other;Other | 0.000 | 0.006 | 0.001 |
| k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Actinomycetaceae;g__Actinomyces | 0.000 | 0.000 | 0.073 |
| k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Corynebacteriaceae;g__Corynebacterium | 0.003 | 0.000 | 0.000 |
| k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Microbacteriaceae;g__ | 0.000 | 0.000 | 0.012 |
| k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Nocardiaceae;g__ | 0.009 | 0.000 | 0.007 |
| k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Nocardiaceae;g__Rhodococcus | 0.000 | 0.000 | 0.001 |
| k__Bacteria;p__Bacteroidetes;c__Sphingobacteriia;o__Sphingobacteriales;f__;g__ | 0.000 | 0.002 | 0.000 |
| k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Chitinophagaceae;g__ | 0.003 | 0.000 | 0.002 |
| k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Chitinophagaceae;g__Sediminibacterium | 0.104 | 0.000 | 0.000 |
| k__Bacteria;p__Chlamydiae;c__Chlamydiia;o__Chlamydiales;f__Rhabdochlamydiaceae;g__Candidatus Rhabdochlamydia | 0.000 | 0.001 | 0.001 |
| k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__Herpetosiphonales;f__;g__ | 0.000 | 0.000 | 0.002 |
| k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Alicyclobacillaceae;g__Alicyclobacillus | 0.000 | 0.253 | 0.000 |
| k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;Other | 0.000 | 0.001 | 0.000 |

| | | | |
|---|---|---|---|
| k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Anoxybacillus | 0.000 | 0.181 | 0.000 |
| k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus | 0.009 | 0.065 | 0.000 |
| k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Paenibacillaceae;g__Paenibacillus | 0.000 | 0.050 | 0.000 |
| k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus | 0.000 | 0.000 | 0.053 |
| k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Catonella | 0.000 | 0.000 | 0.041 |
| k__Bacteria;p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Leptotrichiaceae;g__Leptotrichia | 0.000 | 0.000 | 0.025 |
| k__Bacteria;p__Planctomycetes;c__BD7-11;o__;f__;g__ | 0.000 | 0.000 | 0.037 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae;g__Phenylobacterium | 0.003 | 0.000 | 0.001 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;Other;Other | 0.000 | 0.001 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__;g__ | 0.000 | 0.182 | 0.001 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Bradyrhizobiaceae;g__ | 0.000 | 0.002 | 0.002 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Brucellaceae;g__Ochrobactrum | 0.000 | 0.001 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae;g__ | 0.000 | 0.001 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae;g__Hyphomicrobium | 0.000 | 0.002 | 0.002 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Methylobacteriaceae;g__Methylobacterium | 0.002 | 0.000 | 0.001 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Phyllobacteriaceae;g__Mesorhizobium | 0.028 | 0.000 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Acetobacteraceae;g__Acidocella | 0.000 | 0.000 | 0.031 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Rhodospirillaceae;g__ | 0.001 | 0.000 | 0.002 |
| k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;Other | 0.000 | 0.001 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;Other;Other;Other | 0.000 | 0.000 | 0.002 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;Other;Other | 0.000 | 0.000 | 0.001 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Burkholderia | 0.000 | 0.000 | 0.034 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;Other | 0.000 | 0.000 | 0.001 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Acidovorax | 0.003 | 0.000 | 0.057 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Comamonas | 0.004 | 0.000 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Methylibium | 0.003 | 0.000 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__ | 0.001 | 0.000 | 0.000 |

| | | | |
|---|---|---|---|
| Oxalobacteraceae;Other | | | |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__ | 0.000 | 0.000 | 0.004 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Methylophilales;f__Methylophilaceae;g__ | 0.000 | 0.000 | 0.002 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Neisseriales;f__Neisseriaceae;g__Neisseria | 0.000 | 0.000 | 0.051 |
| k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Procabacteriales;f__Procabacteriaceae;g__ | 0.000 | 0.000 | 0.025 |
| k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__0319-6G20;g__ | 0.000 | 0.001 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter | 0.000 | 0.021 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus | 0.000 | 0.042 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter | 0.259 | 0.008 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;Other | 0.028 | 0.000 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__ | 0.000 | 0.000 | 0.002 |
| k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Sinobacteraceae;g__ | 0.114 | 0.000 | 0.000 |
| k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__ | 0.093 | 0.000 | 0.055 |
| k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Stenotrophomonas | 0.139 | 0.000 | 0.000 |

**Table A.3. Impact of various factors on water chemistry from all Batch 1 samples (ANOSIM, Primer 6).** Euclidean distance matrix was constructed from normalized water chemistry data. "Strata" is based on sampling design, with permutation = 999.

| Factor | Strata | Global R | P |
|---|---|---|---|
| **Utility & Rig & Pipe** | | 0.789 | 0.001 |
| **Utility.rig** | | 0.720 | 0.001 |
| **Utility** | | 0.713 | 0.001 |
| **Rig** | Utility | 0.332 | 0.001 |
| **Pipe Material** | Utility.rig | 0.448 | 0.001 |
| **Stagnation** | Utility.rig | 0.104 | ***0.182*** |

**Table A.4. Normality check of water chemistry data and Kruskal-Wallis analysis results.** Statistical significance was set to 0.05.

| | Shapiro-Wilk normality test | | | Kruskal-Wallis (Utility) | | Kruskal-Wallis (Pipe Material) | | Kruskal-Wallis (Stagnation) | |
|---|---|---|---|---|---|---|---|---|---|
| | **W** | **P** | **Normality (Y/N)** | **Chi-squared** | **P** | **Chi-squared** | **P** | **Chi-squared** | **P** |
| **Temperature** | 0.96 | 2.6e-03 | N | 28.0 | 1.3e-05 | ***5.21*** | ***0.16*** | ***2.48*** | ***0.12*** |
| **pH** | 0.88 | 1.9e-07 | N | 90.6 | < 2.2e-16 | ***0.0094*** | ***1 (0.9998)*** | ***0.001*** | ***0.97*** |
| **Total chlorine** | 0.96 | 2.8e-03 | N | 56.2 | 1.8e-11 | ***6.67*** | ***0.08*** | ***1.57*** | ***0.21*** |
| **Free chlorine[a]** | 0.92 | 1.4e-04 | N | 69.7 | 2.6e-14 | ***8.26*** | ***0.08*** | 4.90 | 0.03 |
| **Turbidity** | 0.25 | <2.2e-16 | N | 20.6 | 3.9e-04 | 8.79 | 0.03 | ***0.97*** | ***0.32*** |
| **Pb** | 0.36 | < 2.2e-16 | N | 19.4 | 6.5e-04 | 51.7 | 3.4e-11 | 10.1 | 0.002 |
| **Cu** | 0.66 | 8.8e-14 | N | 49.5 | 4.5e-10 | 39.0 | 1.7e-08 | 17.0 | 3.82e-05 |
| **Zn** | 0.65 | 3.4e-14 | N | ***7.63*** | ***0.10*** | 36.3 | 6.5e-08 | 3.84 | 0.05 |
| **Na** | 0.82 | 1.5e-09 | N | 95.1 | < 2.2e-16 | ***0.03*** | ***0.999*** | ***0.013*** | ***0.91*** |
| **Mg** | 0.85 | 9.3e-09 | N | 94.1 | < 2.2e-16 | ***0.21*** | ***0.98*** | ***0.10*** | ***0.75*** |
| **Al** | 0.87 | 6.6e-08 | N | 79.7 | < 2.2e-16 | ***1.38*** | ***0.71*** | ***0.57*** | ***0.45*** |
| **Si** | 0.73 | 3.1e-12 | N | 90.8 | < 2.2e-16 | ***0.08*** | ***0.99*** | ***0.05*** | ***0.83*** |
| **P** | 0.83 | 3.3e-09 | N | 92.0 | < 2.2e-16 | ***0.80*** | ***0.85*** | ***0.18*** | ***0.67*** |
| **$NO_3$ as N** | 0.63 | 1.5e-14 | N | 47.8 | 1.1e-09 | ***1.39*** | ***0.71*** | ***1.20*** | ***0.27*** |
| **$SO_4$** | 0.73 | 3.0e-12 | N | 79. 4 | 2.4e-16 | ***0.32*** | ***0.96*** | ***0.28*** | ***0.60*** |
| **TOC as C** | 0.88 | 1.6e-07 | N | 86.6 | < 2.2e-16 | ***1.17*** | ***0.76*** | ***0.57*** | ***0.45*** |
| **K** | 0.69 | 2.5e-13 | N | 94.3 | < 2.2e-16 | ***0.13*** | ***0.99*** | ***0.01*** | ***0.92*** |
| **Ca** | 0.81 | 6.7e-10 | N | 91.4 | < 2.2e-16 | ***0.27*** | ***0.97*** | ***0.11*** | ***0.74*** |
| **Fe** | 0.34 | < 2.2e-16 | N | 47.2 | 1.4e-09 | ***3.36*** | ***0.34*** | ***1.94*** | ***0.16*** |
| **F** | 0.93 | 5.7e-05 | N | 75.5 | 1.6e-15 | ***0.96*** | ***0.81*** | ***0.44*** | ***0.51*** |
| **Cl** | 0.84 | 6.0e-09 | N | 95.1 | < 2.2e-16 | ***0.29*** | ***0.96*** | ***0.11*** | ***0.74*** |

[a] Free chlorine was compared among chlorinated utilities only (A-D): as free chlorine is not meaningful in chloraminated system.

**Table A.5. Pair-wise comparison of microbiome distance matrices across utilities (ANOSIM, permutation=999).**

| Utility Pair | Unweighted UniFrac | | Weighted UniFrac | |
|---|---|---|---|---|
| | Global R | P | Global R | P |
| **A, E** | 0.886 | 0.001 | 0.394 | 0.001 |
| **A, B** | 0.948 | 0.001 | 0.249 | 0.001 |
| **A, C** | 0.997 | 0.001 | 0.883 | 0.001 |
| **A, D** | 0.955 | 0.001 | 0.451 | 0.001 |
| **E, B** | 0.93 | 0.001 | 0.399 | 0.001 |
| **E, C** | 0.997 | 0.001 | 0.757 | 0.001 |
| **E, D** | 0.85 | 0.001 | 0.398 | 0.001 |
| **B, C** | 1 | 0.001 | 0.869 | 0.001 |
| **B, D** | 0.696 | 0.001 | 0.497 | 0.001 |
| **C, D** | 0.999 | 0.001 | 0.822 | 0.001 |

**Table A.6. Impact of various factors on microbiome composition within each utility. Adonis test was applied using weighted UniFrac distance matrix.** Impact size ($R^2$) and significance (P) of four factors at each utility was reported. Data is in the form of $R^2$ value (P value), with significance level set at 0.05.

| Factor | Strata | A | B | C | D | E |
|---|---|---|---|---|---|---|
| **Rig** | | 0.245 (0.001) | 0.176 (0.001) | 0.379 (0.001) | 0.144 (0.001) | 0.176 (0.001) |
| **Pipe material** | Rig | 0.105 (0.002) | 0.097 (0.012) | 0.124 (0.001) | 0.211 (0.001) | 0.324 (0.001) |
| **Stagnation** | Rig | 0.038 (0.023) | 0.035 (***0.054***) | 0.092 (0.001) | 0.085 (0.001) | 0.128 (0.001) |
| **Batch** | Rig.Pipe | 0.039 (0.011) | 0.041 (0.016) | 0.028 (0.010) | 0.026 (0.011) | 0.004 (***0.818***) |

**Table A.7. Microbial Classes that contribute to over 80% similarity of samples within the same utility.**

| Utility | Class | Avg. Abund. | Contrib% | Cum.% |
|---|---|---|---|---|
| **A (69% average similarity)** | *Alphaproteobacteria* | 0.41 | 48 | 48 |
| | *Actinobacteria* | 0.31 | 32 | 80 |
| | *4C0d-2* | 0.08 | 8.2 | 88 |
| **B (76% average similarity)** | *Alphaproteobacteria* | 0.49 | 55 | 55 |
| | *Actinobacteria* | 0.28 | 28 | 83 |
| **C (77% average similarity)** | *Alphaproteobacteria* | 0.72 | 80 | 80 |
| | *Betaproteobacteria* | 0.18 | 16 | 95 |
| **D (66% average similarity)** | *Alphaproteobacteria* | 0.32 | 41 | 41 |
| | *Betaproteobacteria* | 0.19 | 16 | 56 |
| | *Actinobacteria* | 0.15 | 15 | 72 |
| | *Gammaproteobacteria* | 0.15 | 12 | 84 |
| **E (69% average similarity)** | *Alphaproteobacteria* | 0.44 | 52 | 52 |
| | *Betaproteobacteria* | 0.22 | 21 | 74 |
| | *Actinobacteria* | 0.15 | 14 | 88 |

**Table A.8. Top 10 indicator taxa for each utility.**

| Utility | Taxonomy | Indicator value | P | Avg. RA[a] |
|---|---|---|---|---|
| A | k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Staphylococcus | 0.863 | 0.0002 | 9.23E-03 |
| | k__Bacteria;p__Planctomycetes;c__Phycisphaerae;o__Phycisphaerales;f__;g__ | 0.844 | 0.0002 | 3.53E-03 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__;g__ | 0.812 | 0.0002 | 3.08E-02 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Novosphingobium | 0.804 | 0.0002 | 5.93E-03 |
| | k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Alicyclobacillaceae;g__Alicyclobacillus | 0.769 | 0.0002 | 8.43E-03 |
| | k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__ | 0.746 | 0.0002 | 3.06E-03 |
| | k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas | 0.711 | 0.0006 | 1.54E-02 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae;g__Agrobacterium | 0.704 | 0.0004 | 5.89E-03 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Acidovorax | 0.675 | 0.0002 | 6.39E-02 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Hyphomonadaceae;g__ | 0.653 | 0.0002 | 1.64E-02 |
| B | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Hydrogenophaga | 0.787 | 0.0002 | 3.33E-02 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae;g__Mycoplana | 0.636 | 0.0006 | 2.86E-02 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae;g__Caulobacter | 0.585 | 0.0096 | 3.41E-03 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Erythrobacteraceae;g__Erythrobacter | 0.528 | 0.0002 | 6.02E-05 |
| | k__Bacteria;p__Cyanobacteria;c__4C0d-2;o__MLE1-12;f__;g__ | 0.528 | 0.0016 | 1.12E-01 |
| | k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__PSB-M-3 | 0.34 | 0.0002 | 3.55E-05 |
| | k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;Other;Other | 0.313 | 0.0002 | 8.87E-06 |
| | k__Bacteria;p__Firmicutes;c__Clostridia;o__OPB54;f__;g__ | 0.265 | 0.005 | 3.36E-05 |
| | k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__Alteromonadaceae;g__Cellvibrio | 0.262 | 0.0038 | 6.97E-06 |
| | k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Ellin5290;f__;g__ | 0.236 | 0.0078 | 2.53E-05 |

| | | | | |
|---|---|---|---|---|
| | k__Bacteria;p__Verrucomicrobia;c__[Methylacidiphilae];o__Methylacidiphilales;f__LD19;g__ | 0.991 | 0.0002 | 8.44E-03 |
| | k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Methylococcales;f__Methylococcaceae;g__Methylocaldum | 0.991 | 0.0002 | 5.08E-03 |
| | k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales;f__Legionellaceae;g__ | 0.981 | 0.0002 | 6.53E-03 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae;g__ | 0.971 | 0.0002 | 2.92E-03 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Methylophilales;f__Methylophilaceae;g__ | 0.952 | 0.0002 | 2.74E-02 |
| | k__Bacteria;p__TM6;c__;o__;f__;g__ | 0.939 | 0.0002 | 1.47E-03 |
| **C** | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae;g__Hyphomicrobium | 0.91 | 0.0002 | 9.36E-02 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;Other | 0.907 | 0.0002 | 8.82E-04 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Methylophilales;f__Methylophilaceae;Other | 0.903 | 0.0002 | 5.61E-04 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae;Other | 0.870 | 0.0002 | 2.82E-04 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Acetobacteraceae;g__ | 0.857 | 0.0002 | 4.10E-03 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Limnobacter | 0.852 | 0.0002 | 5.94E-03 |
| | k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales;f__Legionellaceae;Other | 0.843 | 0.0002 | 2.92E-04 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__;f__;g__ | 0.839 | 0.0002 | 1.85E-02 |
| **C** | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Burkholderia | 0.834 | 0.0002 | 2.52E-03 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Ralstonia | 0.833 | 0.0002 | 1.18E-01 |
| | k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales;f__Legionellaceae;g__Legionella | 0.828 | 0.0002 | 3.05E-04 |
| | k__Bacteria;p__Chlamydiae;c__Chlamydiia;o__Chlamydiales;f__;g__ | 0.824 | 0.0002 | 5.08E-04 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Blastomonas | 0.811 | 0.0002 | 4.43E-03 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;Other;Other | 0.809 | 0.0002 | 3.95E-04 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__;g__ | 0.796 | 0.0002 | 4.81E-04 |
| D | k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Sinobacteracea | 0.929 | 0.0002 | 1.36E-01 |

| | | | | |
|---|---|---|---|---|
| | e;g__Nevskia | | | |
| | k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__ | 0.889 | 0.0002 | 1.43E-02 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Polaromonas | 0.885 | 0.0002 | 3.47E-04 |
| D | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Sphingobium | 0.799 | 0.0002 | 1.40E-02 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Methylibium | 0.744 | 0.0002 | 1.65E-02 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Acetobacteraceae;g__Roseococcus | 0.713 | 0.0002 | 1.29E-03 |
| | k__Bacteria;p__Acidobacteria;c__Solibacteres;o__Solibacterales;f__[Bryobacteraceae];g__ | 0.713 | 0.0002 | 1.27E-03 |
| | k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Chitinophagaceae;g__ | 0.694 | 0.0002 | 1.33E-02 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__ | 0.675 | 0.0076 | 7.78E-04 |
| | k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Cryomorphaceae;g__Fluviicola | 0.670 | 0.0002 | 8.45E-04 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Nitrosomonadales;f__Nitrosomonadaceae;g__ | 0.998 | 0.0002 | 1.27E-02 |
| | k__Bacteria;p__Planctomycetes;c__Planctomycetia;o__Planctomycetales;f__Planctomycetaceae;g__Planctomyces | 0.976 | 0.0002 | 8.90E-03 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae;g__ | 0.969 | 0.0002 | 3.88E-03 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae;Other | 0.942 | 0.0002 | 2.82E-03 |
| | k__Bacteria;p__Planctomycetes;c__Planctomycetia;o__Pirellulales;f__Pirellulaceae;g__Pirellula | 0.891 | 0.0002 | 4.61E-04 |
| E | k__Bacteria;p__Bacteroidetes;c__Sphingobacteriia;o__Sphingobacteriales;f__;g__ | 0.830 | 0.0002 | 7.99E-03 |
| | k__Bacteria;p__Nitrospirae;c__Nitrospira;o__Nitrospirales;f__Nitrospiraceae;g__Nitrospira | 0.796 | 0.0002 | 1.74E-03 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Methylobacteriaceae;g__Methylobacterium | 0.776 | 0.0002 | 6.78E-02 |
| | k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Rhodocyclales;f__Rhodocyclaceae;g__ | 0.773 | 0.0002 | 1.53E-03 |
| | k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Sphingopyxis | 0.748 | 0.0002 | 5.74E-02 |

[a] Average RA is calculated based on average number of RA within same utility samples using taxonomy Table A.t genus level (rarefied table). Rank is within each utility's samples and omitted "0" ones.

**Table A.9. Association of each variable with the microbiome across all Batch 1 samples. (adonis, permutations = 999, alpha = 0.05, {vegan}, R)**

| Parameter | Unweighted UniFrac | | Weighted UniFrac | |
|---|---|---|---|---|
| | $R^2$ | Pr(>F) | $R^2$ | Pr(>F) |
| **Mg** (ppb) | 0.065 | 0.001 | 0.132 | 0.001 |
| **pH** | 0.075 | 0.001 | 0.131 | 0.001 |
| **F** (mg/L) | 0.051 | 0.001 | 0.118 | 0.001 |
| **P** (ppb) | 0.066 | 0.001 | 0.114 | 0.001 |
| **$SO_4$** (mg/L) | 0.048 | 0.001 | 0.108 | 0.001 |
| **Ca** (ppb) | 0.056 | 0.001 | 0.104 | 0.001 |
| **Free chlorine** (mg/L) | 0.037 | 0.001 | 0.082 | 0.001 |
| **Na** (ppb) | 0.045 | 0.001 | 0.082 | 0.001 |
| **Al** (ppb) | 0.047 | 0.001 | 0.078 | 0.001 |
| **K** (ppb) | 0.037 | 0.001 | 0.069 | 0.001 |
| **Total chlorine** (mg/L) | 0.036 | 0.001 | 0.069 | 0.001 |
| **Cl** (mg/L) | 0.043 | 0.001 | 0.069 | 0.002 |
| **Si** (ppb) | 0.041 | 0.001 | 0.063 | 0.016 |
| **Temperature** (F) | 0.030 | 0.003 | 0.061 | 0.001 |
| **TOC as C** (mg/L) | 0.031 | 0.004 | 0.053 | 0.025 |
| **Cu** (ppb) | 0.028 | 0.001 | 0.039 | 0.022 |
| **Zn** (ppb) | 0.016 | 0.004 | 0.023 | 0.033 |
| **Fe** (ppb) | 0.016 | 0.011 | 0.020 | 0.047 |
| **Pb** (ppb) | 0.016 | 0.008 | 0.017 | ***0.102*** |
| **$NO_3$ as N** (mg/L) | 0.015 | 0.011 | 0.011 | ***0.325*** |
| **Turbidity** (NTU) | 0.011 | ***0.143*** | 0.007 | ***0.632*** |

**Table A.10. Summary of Utility characteristics.**

| Utility | Sampling time | Treatment Process | Disinfectant type |
|---|---|---|---|
| **A** | Nov. 27-29th, 2012 | Conventional treatment plus GAC contactors, UV disinfection as second barrier, chlorine disinfectant, and corrosion inhibitor. | Chlorine |
| **B** | Dec. 12-13th, 2012 | Conventional treatment, chlorine disinfectant and corrosion inhibitor. | Chlorine |
| **C** | Nov. 27-29th, 2012 | Conventional treatment with chlorine disinfection, lime and $CO_2$ added for water stability, no corrosion inhibitor. | Chlorine |
| **D** | Dec. 17-18th, 2012 | Conventional treatment plus ultra-filtration membranes, chlorine disinfection and corrosion inhibitor. | Chlorine |
| **E** | Jan. 13-16th, 2013 | Conventional treatment with UV disinfection and chloramines disinfection, no corrosion inhibitor, high alkalinity. | Chloramine |

# CHAPTER 3. IMPACT OF WATER HEATER TEMPERATURE SETTING AND WATER USE FREQUENCY ON THE BUILDING PLUMBING MICROBIOME

Pan Ji, William J. Rhoads, Marc A. Edwards, Amy Pruden

## ABSTRACT

Hot water plumbing is an important conduit of microbes into the indoor environment and can increase risk of opportunistic pathogens (e.g., *Legionella pneumophila*). We examined the combined effects of water heater temperature (39, 42, 48, 51, and 58°C), pipe orientation (upward/downward), and water use frequency (21, 3, and 1 flush/week) on the microbial composition at the tap using a pilot-scale pipe rig. 16S rRNA gene amplicon sequencing indicated that bulk water and corresponding biofilm typically had distinct taxonomic compositions ($R^2_{Adonis}$ = 0.246, $P_{Adonis}$ = 0.001), yet highly similar predicted functions based on PICRUSt analysis ($R^2_{Adonis}$ = 0.013, $P_{Adonis}$ = 0.014). While a prior study had identified 51°C under low water use frequency to enrich *Legionella* at the tap, here we reveal that 51°C is also a threshold above which there are marked effects of the combined influences of temperature, pipe orientation, and use frequency on taxonomic and functional composition. A positive association was noted between relative abundances of *Legionella* and mitochondrial DNA of *Vermamoeba*, a genus of amoebae that can enhance amplification and virulence of some pathogens. This study takes a step towards intentional control of the plumbing microbiome and highlights the importance of microbial ecology in governing pathogen proliferation.

## INTRODUCTION

There is growing interest in the microbiome inhabiting the built environment and its potential implications to human health (Hanski et al., 2012; Oberauner et al., 2013; Kelley and Gilbert, 2013). Building plumbing systems deliver water, along with its resident microbes, into homes and other buildings. Most attention has been directed towards opportunistic pathogens (OPs), such as *Legionella pneumophila*, which are particularly challenging to control as they are native to the fresh water environment. Collectively, OPs residing in building plumbing have become the top contributor of waterborne disease in the U.S. and other developed countries (Ashbolt, 2015; Beer et al., 2015; Hilborn et al., 2013), imposing an economic burden over $1 billion/ year in the U.S. alone (Falkinham III et al., 2015). Thus, understanding the composition of the microbiome at the tap, and its contribution of pathogens and general microbiota to built environment and human microbiomes is of great importance.

Hot water plumbing systems are of special interest as a typical route of OPs exposure is inhalation of aerosols and/or direct contact during showering (Feazel et al., 2009). Elevating the temperature of hot water storage tanks is known to be an effective measure to control *L. pneumophila* (Falkinham III et al., 2015; Buse and Ashbolt, 2011; ANSI/ASHARE, 2015; Darelid et al., 2002), and 60°C has been demarcated as a regulatory threshold (VHA Directive 1061; Department of Veterans Affairs, 2014). However, motives to reduce energy usage and avoid scalding are in conflict with higher temperature settings (Lévesque et al., 2004; Rhoads et al., 2016a) and long-term effectiveness of thermal eradication has been questioned (Farhat et al. 2012). In contrast to a number of studies directly focusing on *Legionella* in hot water systems (e.g., Bargellini et al., 2011; Borella et al., 2004; Zacheus and Martikainen, 2004; Zietz et al., 2001; Mathys et al., 2008), very few have more broadly examined the microbiome as a whole. Prior studies included field surveys on microbial communities collected from shower head/hose biofilms (Feazel et al., 2015; Proctor et al., 2016) and bulk water samples from a hospital hot water system, where the effectiveness of monochloramine disinfection was examined (Baron et al., 2014). However, a systematic examination of the combined effects of water heater temperature and system design/operational conditions on the actual microbiome composition at the tap has been lacking.

While hot water storage tanks may remain at elevated temperature settings, the reality at the distal tap is that the biofilm experiences only brief heat shock. Water use frequency at a given tap is thus of particular interest, as it dictates the frequency of heat shock as well as the length of time for the biofilm microbiota to recover from the perturbation, which could be further influenced by pipe orientation and associated convective mixing conditions (Rhoads et al., 2016b). In the bulk water phase, in-building stagnation was found to have a marked effect on the microbial composition of cold water lines (Ji et al., 2015; Lautenschlager et al., 2010). Moreover, warmer indoor temperature together with overnight stagnation could increase microbial metabolic activity level and alter the functional profile (Zhang et al., 2015). We thus hypothesized that different water use frequencies would alter the microbiome composition of both the biofilm and bulk water.

The relationship between the microbial composition of the biofilm and bulk water is also widely debated and likely to be influenced by operating conditions. Biofilm is often perceived to be a reservoir for OPs, where protist hosts can enhance their amplification and, along with the biofilm matrix, protect them from unfavorable conditions such as chlorine/monochloramine disinfection (Cooper and Hanlon, 2010; Pryor et al., 2004) and low nutrients (Lau and Ashbolt, 2009). However, microbes entrained in the biofilm must be released into the bulk water for relevance to human exposure, which can occur long after the initial entrapment/attachment of the microbe to the biofilm (Howe et al., 2002). Interestingly, OPs in the two phases do not necessarily respond the same way under a given condition. For example, at room temperature, stagnation predisposed bulk water to *Legionella* proliferation (Wang et al., 2012), but resulted in lowest *Legionella* concentrations in biofilm relative to those under turbulent and laminar flow conditions (Liu et al., 2006). Such questions call for a mechanistic understanding of the factors shaping the biofilm and bulk water microbiomes of domestic hot water systems.

To overcome complexities encountered in field studies, while maintaining relevance to real-world conditions, we constructed two identical recirculating hot water plumbing rigs with distal taps operated in triplicate under three water use scenarios (21, 3, and 1 flush/week) and two orientations (upward and downward flow). Water heater temperature was elevated in a step-wise

fashion from 39 to 58 °C, while an identical control rig remained at 39°C throughout the 15-month study period. Illumina amplicon sequencing of 16S rRNA genes was applied to map the bacterial and archaeal community composition of influent, recirculating line and distal taps with paired biofilm and first-flush bulk water samples. Metagenomic data was predicted using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt 1.0.0) to characterize hypothetical functionalities of resident microbes. Overall the study takes a step towards identifying to what extent building owners/operators may be able to intentionally control the microbiome composition at the tap through practical measures, and addresses a critical knowledge gap at the intersection of water infrastructure, energy saving and public health.

## MATERIALS AND METHODS

**System description.** Two identical lab-scale hot water plumbing rigs were designed for paired comparison of the impact of water heater temperature and use frequency (Fig. B.1), as reported previously in studies focused on the occurrence of *Legionella* (Rhoads et al., 2015; Rhoads et al., 2016b). After 5- month acclimation at 39°C, the control rig was maintained at 39°C, while the experimental rig was increased from 39 to 42, 48, 51 and 58°C at intervals of 2-3 months. Both rigs were fed with Blacksburg (VA, USA) tap water pre-treated with Pentek GAC-BB granular activated carbon cartridges (Pentair Residential Filtration, Milwaukee, WI, USA). Factors examined within each rig included: water heater temperature setting (°C), time (5, 8, 10, 13, and 15 months), sample location (influent, recirculating line, and distal tap), system (control rig and experimental rig), pipe orientation (upward and downward flow), phase (biofilm and bulk water), and water use frequency (high, medium, low: 21, 3, 1 flush/week respectively) in triplicate distal pipes. Notably, distal taps achieved room temperature within 25 minutes after flushing, in contrast to the recirculating line, which was consistently maintained at the temperature setting (Rhoads et al., 2015). If not specified, “temperature” denotes the water heater setting, while “effective” temperature indicates the average in the distal pipe during stagnation (i.e., across the flush/stagnation cycle).

**Sample collection and processing.** Biofilm swab (~65 $cm^2$) and first-flush bulk water (~500 mL) samples were collected after 2- to 3-month acclimation time at each condition, at the end of the stagnation cycle. Bulk water samples were concentrated onto 0.22 μm pore-size sterile mixed cellulose ester filters (Millipore, Billerica, MA, USA). Filters and swabs were transferred to lysis tubes and subject to DNA extraction using FastDNA Spin Kits (MP Biomedicals, Solon, OH, USA) according to manufacturer protocol. Illumina 16S rRNA amplicon sequencing was performed following the Earth Microbiome Project 16S rRNA Amplification Protocol Version 4_13 (http://www.earthmicrobiome.org/emp-standard-protocols/16s/). Minor changes were as previously described (Ji et al., 2015), including using Molecular Biology Grade Water (Quality Biological, Gaithersburg, MD, USA) and pooling 200 ng of amplicon from each sample before clean-up using QIAquick PCR Purification Kit (QIAGEN, Valencia, CA, USA). Universal primers 515f and barcoded 806r were used for PCR amplification. Illumina 16S rRNA amplicon sequencing (paired-end 250 bp, MiSeq platform) was performed by the Virginia Bioinformatics Institute (Blacksburg, VA, USA).

**Microbiome data analyses.** Paired-end reads were stitched using PANDAseq (Masella et al., 2012) and filtered based on quality score (> 0.90, range 0-1) and sequence length (252-255 bp) following the EDAMAME course materials (https://edamame-course.github.io/, accessed 01/24/2016). A *de novo* Operational Taxonomic Unit (OTU) picking strategy was adopted in QIIME 1.8.0 (Caporaso et al., 2010) as previously described (Ji et al., 2015). Mitochondrial and chlorophyll sequences were removed from both the OTU Table and the phylogenetic tree before downstream analysis. After removal of singleton OTUs, the sequencing depth of the 427 samples ranged from 25 266 to 304 049 with a median value of 82 078, which translated into 62 494 OTUs in total. All samples were rarefied to the smallest sequencing depth (i.e., 25 266) prior to downstream alpha diversity (in terms of Shannon index, rarefied 10 times) and beta diversity (in terms of weighted UniFrac distance matrix, rarefied 100 times) analyses in QIIME 1.8.0 (Caporaso et al., 2010). Principal Coordinate Analysis was further applied to the beta diversity and visualized in EMPeror (Vázquez-Baeza et al., 2013).

Metagenomic data was predicted from the OTU table using PICRUSt 1.0.0 (Langille *et al.*, 2013). Bray-Curtis distance matrices were calculated from predicted functional profile at Kyoto

Encyclopedia of Genes and Genomes (KEGG) level 3 ({labdsv} version 1.8-0) in R studio (R version 3.2.2, R Core Team 2015). All factors were treated as non-numeric variables, and the relative contribution of each factor on the microbial phylogenetic and functional dissimilarities across samples were determined through Permutational Multivariate Analysis of Variance (Adonis, Anderson, 2001) among *a priori* defined groups in R studio (permutation = 999, {vegan} version 2.3-0, Oksanen et al., 2016). Adonis is less sensitive to dispersion (i.e., within-group variation) than alternative methods including ANOSIM (Oksanen et al., 2016). A complementary multivariate homogeneity of groups dispersions (variances) analysis (betadisper; Anderson, 2006) was applied to account for confounding dispersion effect.

To explore the relationship between biofilm and bulk water microbiota, we first examined if both phases trended similarly in response to various factors through Mantel and Procrustes tests (QIIME 1.8.0, Caporaso et al., 2010). Second, we explored to what extent biofilm and bulk water shared microbes via core OTU analysis on both sample groups and individual paired samples (169 pairs of bulk water- biofilm distal tap samples). Based on Levene's test of homogeneity of variances, 1-way ANOVA with Tukey's Honest Significant Differences (HSD) method for pair-wise comparisons, or Kruskal and Wallis test with Tukey and Kramer (Nemenyi) test was applied ({car} version 2.1-2, Fox and Weisberg, 2011; {stat} version 3.2-2; {PMCMR} version 4.1, Pohlert, 2014). A Venn Diagram ({VennDiagram} version 1.6.16, Chen and Boutros, 2011) was constructed for the former. A complementary indicator microbe analysis ({indicspecies} version 1.7.5, De Caceres and Legendre, 2009) was performed to inform the identity of the shared/exclusive OTUs of interest with cutoff indicator values set at 0.7 (both Quantities A and B).

Associations between *Legionella* spp. and other microbial genera were characterized using Maximal Information-based Nonparametric Exploration (MINE) analysis ({minerva} version 1.4.1, Reshef et al., 2011) with a false detection rate (FDR) correction at 0.05 (Benjamini and Hochberg, 1995). Mitochondrial sequences belonging to *Vermamoeba vermiformis* were detected, and the "relative abundance" of *V. vermiformis* was calculated as the ratio of the number of mitochondrial sequences to the number of total non-mitochondrial sequences, which was later incorporated into the taxonomy table the genus-level before MINE analysis.

**Data sharing.** All sequences and metadata for this research are retrievable from QIITA Study ID 10502.

## RESULTS

**Drivers of phylogenetic and functional microbiome dissimilarity.** Phase (biofilm versus bulk water) was the major driver of phylogenetic dissimilarity across all samples ($R^2_{Adonis}$ = 0.246, $P_{Adonis}$ = 0.001), followed by water heater temperature setting ($R^2_{Adonis}$ = 0.172, $P_{Adonis}$ = 0.001), sample location (influent, recirculating line, and distal tap; $R^2_{Adonis}$ = 0.124, $P_{Adonis}$ = 0.001), and time ($R^2_{Adonis}$ = 0.100, $P_{Adonis}$ = 0.001) (Table 3.1). Water use frequency ($R^2_{Adonis}$ = 0.039, $P_{Adonis}$ = 0.001) and pipe orientation ($R^2_{Adonis}$ = 0.013, $P_{Adonis}$ = 0.001) had less pronounced effects on phylogenetic dissimilarity, although their impacts were apparent under certain conditions (e.g., biofilm at T = 58°C, Fig. B.2). At the tap, elevated temperature and time together conferred nearly double the magnitude of impact on phylogenetic microbiome dissimilarity (experimental rig; $R^2_{Adonis}$ = 0.249, $P_{Adonis}$ = 0.001) relative to that of constant temperature and time (control rig; $R^2_{Adonis}$ = 0.112, $P_{Adonis}$ = 0.001), with the most striking contrast between the two rigs at 15 months when the experimental rig was at 58°C (Fig. B.3). Temperature appears to have exerted an influence on the microbiome composition over and above that of succession with time given that: (1) there was no significant difference in the phylogenetic composition between the two rigs during the initial 5-month baseline at 39°C ($P_{Adonis}$ = 0.187) and (2) influent samples did not exhibit a statistically significant temporal trend ($R^2_{Adonis}$ = 0.160, $P_{Adonis}$ = 0.679). Notably, no significant influence of time or temperature was observed in recirculating line samples (control rig, $P_{Adonis}$ = 0.811; experimental rig, $P_{Adonis}$ = 0.119), indicating that effects of time and temperature primarily manifested at the tap.

**Table 3.1. Relative impact of various factors on phylogenetic and functional microbiome dissimilarity across samples** (Adonis and Betadisp, permutation = 999). Both time (t) and water heater temperature setting (T) were treated as non-numeric factors. Statistical significance was set at 0.05: italic grey font indicates non-statistically significant; underlined indicates statistically significant Betadisp result.

| Factor | Phylogenetic | | | Functional | | | Samples |
|---|---|---|---|---|---|---|---|
| | $R^2_{Adonis}$ | $P_{Adonis}$ | $P_{Betadisp}$ | $R^2_{Adonis}$ | $P_{Adonis}$ | $P_{Betadisp}$ | |
| Phase | 0.246 | 0.001 | 0.001 | 0.013 | 0.014 | 0.001 | All samples (n=427) |
| Sample location | 0.124 | 0.001 | 0.001 | 0.032 | 0.002 | 0.044 | All samples (n=427) |
| t | 0.100 | 0.001 | 0.001 | 0.570 | 0.001 | 0.005 | All samples (n=427) |
| T | 0.172 | 0.001 | 0.008 | 0.233 | 0.001 | 0.001 | All samples (n=427), influent was set at 20°C. |
| Water use frequency | 0.039 | 0.001 | 0.042 | *0.005* | *0.443* | *0.042* | Distal tap samples (n=348) |
| Pipe orientation | 0.013 | 0.004 | *0.730* | *0.004* | *0.189* | *0.745* | Distal tap samples (n=348) |
| Phase | 0.311 | 0.001 | 0.001 | 0.021 | 0.007 | 0.001 | Distal tap samples (n=348) |
| Rig | *0.021* | *0.187* | *0.434* | *0.021* | *0.235* | *0.289* | Distal tap at 5 months, 39°C (n = 70) |
| t (@ 39°C) | 0.112 | 0.001 | 0.038 | 0.552 | 0.001 | *0.057* | Control rig distal tap samples (n=175) |
| Phase | 0.466 | 0.001 | 0.001 | 0.032 | 0.016 | 0.022 | Control rig distal tap samples (n=175) |
| t (elevated T) | 0.249 | 0.001 | 0.018 | 0.670 | 0.001 | 0.049 | Experimental rig distal tap samples (n=173) |
| Phase | 0.258 | 0.001 | 0.001 | *0.013* | *0.119* | 0.016 | Experimental rig distal tap samples (n=173) |
| Phase | *0.020* | *0.423* | *0.129* | *0.015* | *0.380* | *0.125* | Recirc. Samples (n=45) |
| Rig | 0.129 | 0.001 | *0.725* | 0.092 | 0.020 | *0.381* | Recirc. Samples (n=45) |
| t (@ 39°C) | *0.110* | *0.811* | *0.081* | 0.757 | 0.001 | *0.645* | Control rig recirc samples (n=27) |
| t (elevated T) | *0.198* | *0.119* | 0.003 | 0.554 | 0.002 | *0.549* | Experimental rig recirc samples (n=28) |
| t | *0.160* | *0.679* | *0.118* | 0.758 | 0.001 | *0.607* | Influent (n=24) |
| Rig | 0.041 | 0.001 | 0.035 | *0.001* | *0.517* | *0.280* | Distal tap samples (n=348) |
| T (below and above) | 0.17543 | 0.001 | 0.001 | 0.20344 | 0.001 | 0.001 | Experimental rig distal tap samples (n=173) |

When considering the microbiome composition from a functional standpoint, employing PICRUSt, there was no significant dissimilarity as a function of phase or sample location (Table 3.1). Instead, temperature and time were most influential, though the magnitude of impact of elevated temperature on function was not as great as that which had been observed on phylogenetic profile (Table 3.1). For the influent water, the functional microbiome was found to be dissimilar with time ($R^2_{Adonis} = 0.758$, $P_{Adonis} = 0.001$).

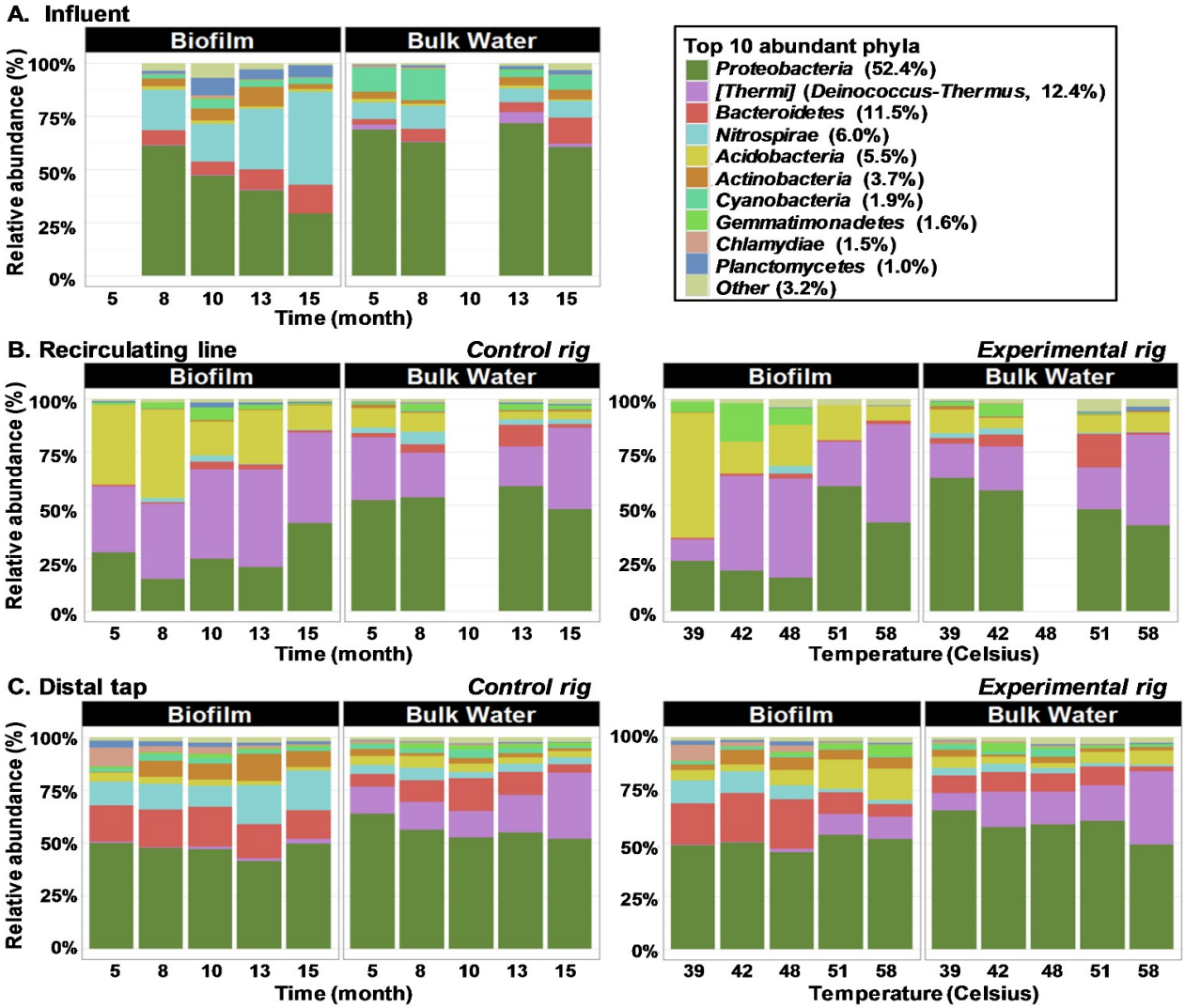


**Fig. 3.1. Taxonomic profiles of sample groups at the phylum level.** Temperature indicates water heater temperature setting. Top abundant phyla were defined as those that were > 1% of total sequences across all samples combined (percentage shown in brackets); the remainder were combined and lumped in a category designated as "Other". The three blank columns correspond to lost samples.

**Effect of combined factors on microbiome composition.** Grouping samples by phase, sample location, temperature and time resulted in 46 sample groups, which shared the same top abundant phyla (i.e., phyla with over 1% of total sequences across all samples), but differed in their relative abundances (Fig. 3.1). For instance, *Nitrospirae* favored the biofilm phase and were enriched with time in the influent and control rig distal taps, but diminished in the experimental rig distal taps with elevated temperature. In the recirculating line, *Nitrospirae* remained at a comparatively low relative abundance, with no obvious temporal trend in either rig. Almost all *Nitrospirae* sequences were classified as *Nitrospira* spp., which are nitrite-oxidizing bacteria with some members recently reported to be "comammox" (i.e., complete ammonia oxidizer) (Daims *et al.*, 2015) and to have an optimal temperature range of 35-40°C. In the recirculating line, higher temperatures (51 and 58°C) were associated with smaller net decrease in ammonia concentration (Table B.1) and concurrently lower relative abundance of *Nitrosomonas* spp. in bulk water (Table B.2), which likely suggests inhibited ammonia oxidization given that influent ammonia level remained stable. Some other ammonia-oxidizing bacteria containing genera, including *Nitrosospira* (Burrell et al., 2001), were not found in the recirculating lines.

**Biofilm versus bulk water: indicator OTUs.** The indicator OTU concept provided a framework for examining the preference of individual microbes for the biofilm versus the bulk water. Ideally, an indicator OTU should meet two criteria: (1) 100% of the OTU sequences belong to the samples from the defined group (*specificity* or *positive predictive value*, indicator value index A = 1.0) and (2) the OTU is detected in 100% of the samples from the defined group (*fidelity* or *sensitivity*, indicator value index B = 1.0) (Cáceres and Lengendre, 2009; Cáceres et al., 2010). Applying 0.7 as the cutoff value for both indicator value indices, we observed that bulk water harbored more indicator OTUs (232) than did biofilm (147). Notably, OTUs within the same genus did not universally serve as indicators for the same phase or sample group. For instance, bulk water phase indicators included 11/631 *Mycobacterium* spp., 34/635 *Meiothermus* spp., and 1/188 *Legionella* spp. OTUs. Similarly, different OTUs within the same genus sometimes indicated different phases. Out of 1 056 OTUs belonging to *Nitrospira* spp., 11 served as indicators for the biofilm phase while 7 others were indicators for the bulk water phase.

This suggests that it may be inappropriate to assume that members of the same genus uniformly serve as indictors in domestic hot water systems.

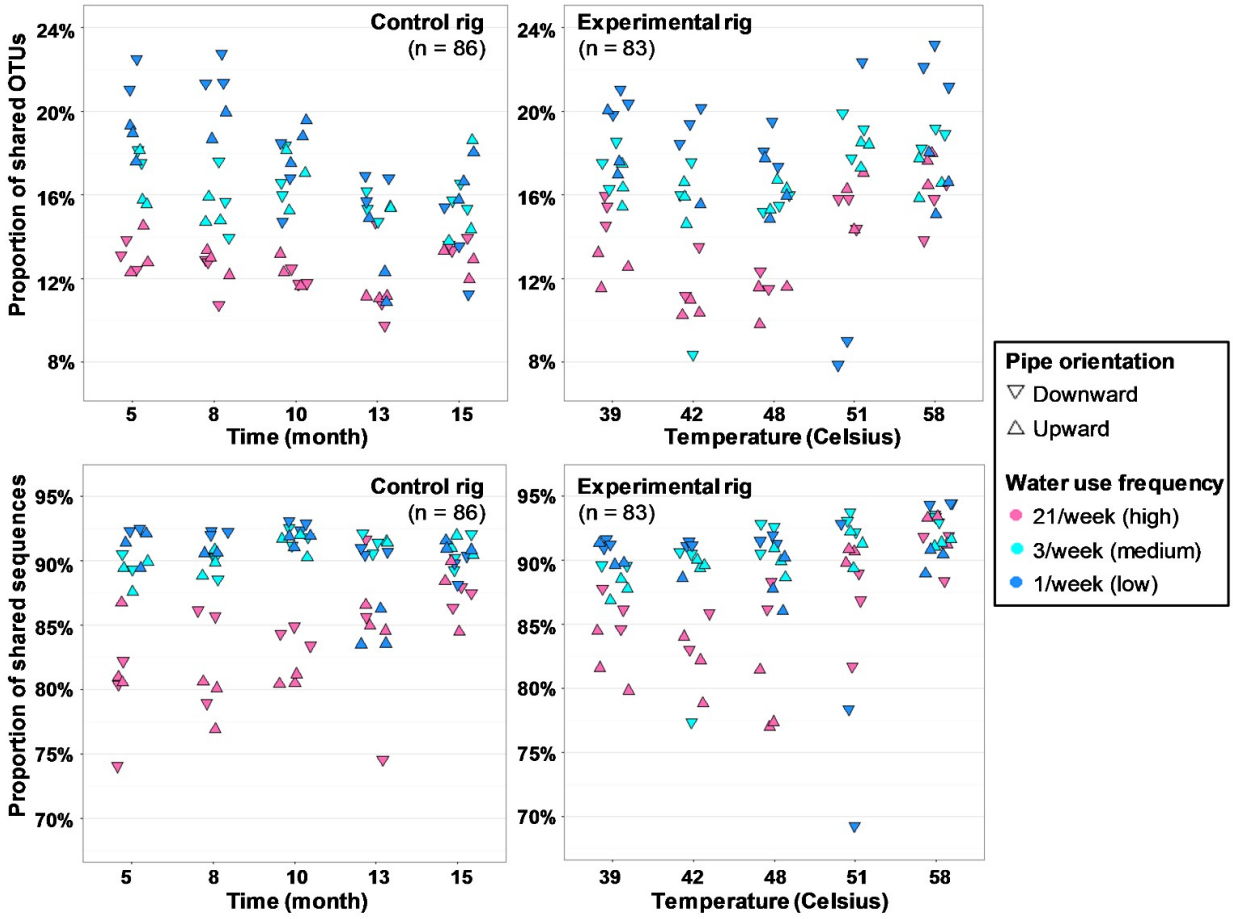


**Fig. 3.2. Proportion of OTUs and sequences shared in paired bulk water and biofilm samples at distal taps.** Proportion was calculated as either number of shared OTUs divided by number of total OTUs between the paired samples (membership); or the number of sequences belonging to shared OTUs divided by the number of total sequences between the paired samples (relative abundance). Color codes indicate water use frequency and shape indicates pipe orientation.

Bulk water and biofilm trended with each other in response to various factors including time and water heater temperature (Mantel: $R^2 = 0.608$, $P = 0.001$; Procrustes $M^2 = 0.698$). In general, each bulk water and biofilm sample pair shared 8-24% OTUs, which comprised 68-95% of the total sequences between the pairs (Fig. 3.2). Comparison between the two rigs suggested that

elevated temperature (T ≥ 51°C) induced a subtle increase in the proportion of shared OTUs between the biofilm and bulk water (Tukey HSD; adjusted $P_{51\text{-}39°C}$ = 0.046, adjusted $P_{58\text{-}39°C}$ = 0.005), although this trend was not significant when viewed as the proportion of shared sequences (adjusted $P_{51\text{-}39°C}$ = 0.992, adjusted $P_{58\text{-}39°C}$ = 0.258) (Table B.3). At T < 51°C, high water use frequency was associated with the fewest shared OTUs and sequences between paired bulk water and biofilm samples (Fig. 3.2; adjusted P < 0.001 for 21/week – 1/week and 3/week – 1/week, Table B.4; Table B.5). Pipe orientation (upward and downward flow) had impact on the relationship between bulk water and biofilm in some situations. For instance, under high water use frequency (21/week), downward pipe harbored higher shared OTUs between phases, except at 10 and 15 months in the control rig (39°C).

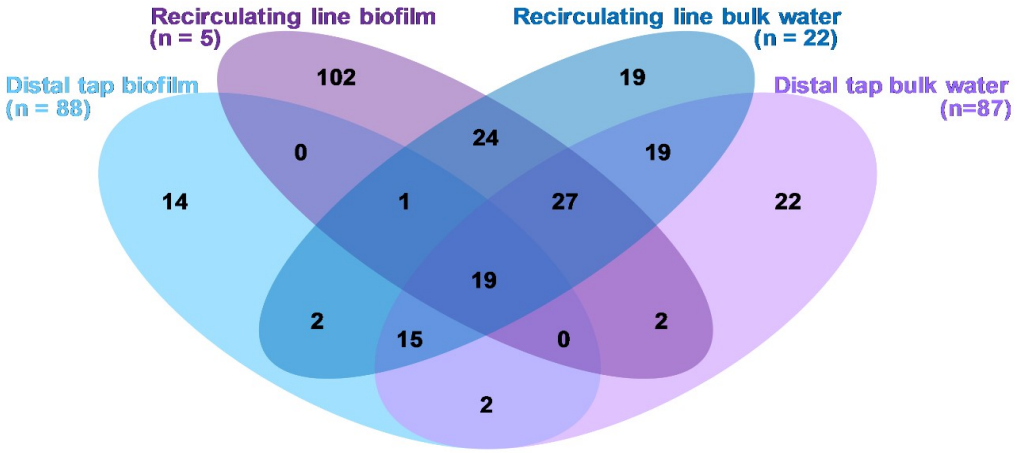


**Fig. 3.3. Venn Diagram of core Operational Taxonomic Units (OTUs) among each of the recirculating line and distal tap phases in the control rig across all sampling dates.**

**Influence of recirculating line microbes on those at the distal taps: core OTUs.** Beta diversity analysis indicated that the greatest contrast occurred between the influent and recirculating line, with the distal taps displaying noticeable overlap with both of these locations (Fig. B.4). Of particular interest was the relative contribution of microbes from the influent and recirculating line to those observed at the tap, which can be conceptualized as survival of allochthonous microbes to the periodically heat-shocked distal tap environment. To facilitate this comparison, control rig samples were subject to core OTU analysis (i.e., OTUs shared among all samples within the given group) corresponding to 4 groups including distal tap biofilm, recirculating line biofilm, distal tap bulk water, and recirculating line bulk water (Fig.

3.3). There were 14, 102, 19, and 22 core OTUs confined to each of these four individual groups, respectively, with a total of 19 core OTUs common to all groups. The vast majority of distal tap bulk water core OTUs (82/106 OTUs) were shared with recirculating line core OTUs (both phases included). Of these, 34 OTUs were only shared with recirculating line bulk water and 2 were only shared with the recirculating line biofilm. Interestingly, 20 core OTUs that were shared between the two biofilm locations were also encompassed in recirculating line bulk water core OTUs. This suggests that significant migration occurs from the recirculating line to the tap via physical transport of bulk water. Biofilm at the tap is also a likely source of bulk water OTUs at the tap, with 36 core OTUs shared between these two phases and 2 of which confined to the tap.

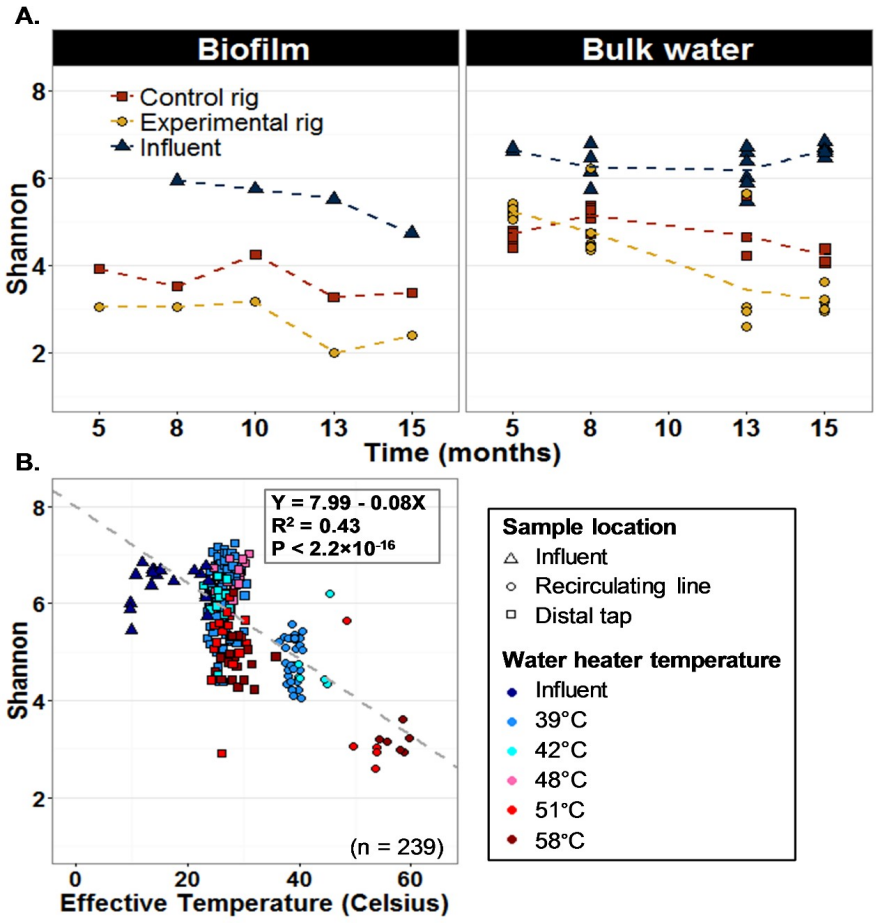


**Fig. 3.4. A. Alpha diversity (Shannon Index) across influent and recirculating lines in both rigs** (note for experimental rig temperature was elevated with time); **B. negative association**

**between alpha diversity and "effective" temperature in bulk water samples** (grey line is linear correlation).

**Negative association between alpha diversity (Shannon Index) and temperature.** Alpha diversity trends were similar for both the biofilm and bulk water phase in that the influent was characterized by the highest Shannon Index, followed by the control rig recirculating line, and the experimental rig recirculating line (Fig. 3.4.A). In the biofilm phase, the experimental rig trended with the control rig with time; whereas in the bulk water phase, the experimental rig experienced a more drastic decrease concurrent with the elevated temperature. Linear regression ({stats} version 3.2.2, R Studio) further confirmed a negative association between Shannon Index and "effective" temperature (i.e., average temperature across the flush/stagnation cycle) in bulk water samples (Fig. 3.4.B). Despite this negative association, influent samples, with "effective" temperatures below 15°C, had lower alpha diversity than the fitted model, which contributed to a peak alpha diversity region around 20 to 25°C. Distal tap samples did not vary much from each other in terms of "effective" temperature (26.8 ± 2.2°C), but displayed a wide range of alpha diversities. Generally, higher temperature ($T \geq 51$°C) appeared to be conducive to lower alpha diversity.

**Patterns in *Legionella* occurrence and association with other microbes.** The relative abundance of *Legionella* spp. decreased with elevated temperature (Fig. 3.5.A), whereas it increased in the bulk water phase of the control rig with time (Fig. 3.5.B). Nevertheless, at T = 51°C, 3 out of 5 low water use frequency (1/week) bulk water samples displayed markedly high relative abundance of *Legionella* spp., 2 of which were from upward oriented pipes. Rhoads et al. (2015) also confirmed that this combination of 51°C and low water use frequency uniquely selected for *L. pneumophila* at distal taps using quantitative PCR. Pipe orientation (upward and downward) alone did not show an obvious impact on enrichment of *Legionella* spp. (Fig. B.5, paired upward/downward pipes).

Mitochondrial sequences belonging to *V. vermiformis*, which are known to serve as hosts for amplification and enhanced virulence of *L. pneumophila* and other pathogens, were detected and noted to trend with *Legionella* spp. in terms of decreased relative abundance with elevated

temperature (Fig. 3.5.C). *V. vermiformis* also tended to be enriched in the bulk water microbiome with time, albeit much slower than *Legionella* spp. (Fig. 3.5.D). MINE analysis identified a total of 143 statistically significant associations between *Legionella* spp. and other microbial taxa under a False Detection Rate (FDR) of 0.05. The most strongly associated genus was mitochondrial *Vermamoeba* spp. (*V. vermiformis*), with Maximal Information Coefficient (MIC) of 0.44. *Legionella* spp. was also found to be positively associated with *Mycobacterium* spp., *Meiothermus* spp., and *Xanthomonadaceae* family (MIC = 0.35, 0.41, and 0.51, respectively); while negatively associated with *Nitrospira* spp. (MIC = 0.28).

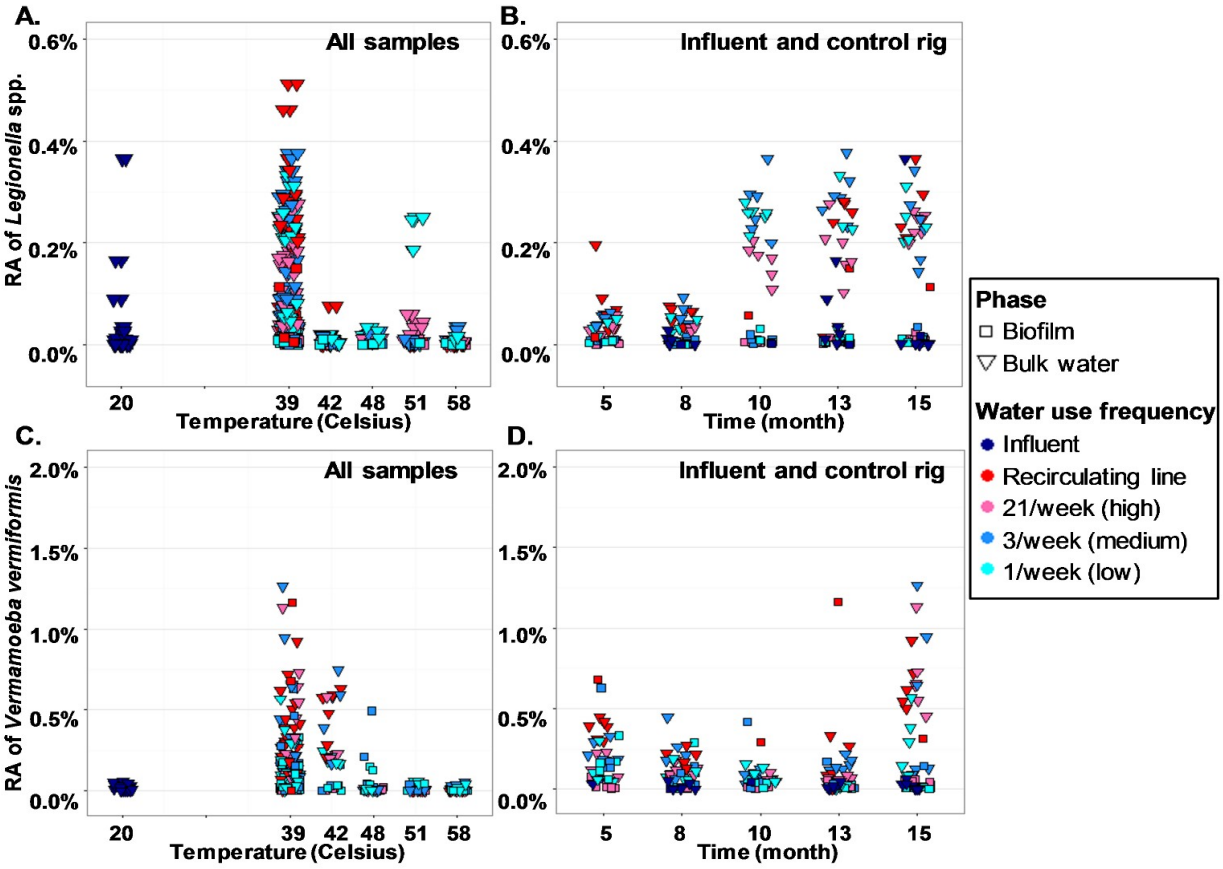


**Fig. 3.5. Relative abundance (RA) of *Legionella* spp.** (top) **and *Vermamoeba vermiformis*** (mitochondrial, bottom) **in all samples under different water heater temperature settings** (left) **or influent and control rig samples across time** (right)**.** Relative abundance was calculated using Illumina 16S rRNA gene amplicon sequencing. Water use frequency and phase are indicated by color and shape, respectively. All influent samples were assigned a uniform temperature of 20°C to facilitate comparison.

## DISCUSSION

This study utilized pilot-scale, replicated pipe rigs to unravel the complexity inherent in building plumbing systems and gain insight into the impact of water heater temperature and water use frequency on the microbiome composition at the tap. The findings have important implications for informing selection of water heater settings as well as water system design, which are currently the targets for new design and optimization strategies for conserving energy and water. Such designs could inadvertently shape the numbers and kinds of microbes entering the built environment via tap water.

A total of 62 494 OTUs were identified across this study, about a quarter of which were accounted for by influent cold tap water samples (15 478 OTUs both bulk water and biofilm). The total number of OTUs identified was about 10 times the highest number reported in prior drinking water microbiome surveys (6 488 in El-Chakhtoura et al., 2015; Lautenschlager et al., 2013; Pinto et al., 2014; Lührig et al., 2015; Proctor et al., 2015), which have focused exclusively on cold water. We found that elevated water heater temperature drove significant change with respect to both the phylogenetic composition and predicted functions of the microbiota at the tap. Interestingly, the bulk water composition shifted with a threshold response around 51°C, whereas the biofilm displayed more of a continuous succession pattern as the temperature was gradually elevated. The findings are consistent with the expectation of the environment selecting microbial populations (the niche theory, Hutchinson, 1957), which could theoretically provide the basis for a prebiotic framework for selection of "beneficial" microbes in the built environment. In interpreting the results throughout, it is important to bear in mind that the analysis applied in this study was entirely DNA-based, and therefore cannot directly distinguish live, dead, or viable but non-culturable microbes. However, systematic comparison across various interconnected locations of the pipe rigs through various operating conditions and time provided an indication of which microbes were likely most adaptive to each location/condition.

While there appeared to be a sharp bifurcation between biofilm and bulk water microbial composition, as reported by others using culture and 16S rRNA gene based clone libraries (Martiny et al., 2005), highly similar functions between the two phases were predicted. This implies functional redundancy of the microbiota, or what has been dubbed as a "portfolio effect" in which positive and negative changes in individual taxa observed in the bulk water and biofilm essentially sum to zero (Allison and Martiny, 2008). Conceptually, the distinct dispersal limitations between planktonic (bulk water) and sessile (biofilm) states would allow for the development of functionally-equivalent, yet taxonomically diverse, species adapting to a similar set of environmental conditions (the neutral theory, Hubbell, 2001). In fact, neutral theory has been proposed in conceptualizing biofilms as microbial "landscapes" (Battin et al., 2007). Rather than a niche-neutral dichotomy, recent studies have recognized the importance of both selective and neutral processes in shaping gastrointestinal (Jeraldo et al., 2012), soil (Dumbrell et al., 2010) and wastewater (Ofiţeru et al., 2010) microbiomes.

Fundamental to this study is the fact that the distal tap environment is distinct from that of the recirculating line. Distal tap water samples essentially consisted of recirculating line inflow that had stagnated and cooled to room temperature (within ~25 minutes) since the previous flush of hot water. Stagnation has been reported to be a significant driving force for shift in potable cold water microbial communities (Lautenschlager et al., 2010; Ji et al., 2015). Within the control rig (T = 39°C), the substantial proportion of core OTUs unique to the distal tap bulk water (Figure 3) suggests that stagnation as short as 8 hours (high water use frequency 21/week) is sufficient for a shift away from the core microbes defining the recirculating line bulk water or biofilm. It remains unclear, though, at what rate conversion from core to satellite microbiome generally occurs and whether elevated water heater temperature would influence the conversion time. The difference observed in as little as 8 hours also demonstrated that the DNA-based approach applied in this study was able to capture microbial community shifts occurring on a short time scale.

The interrelationship between biofilm and bulk water microbes in building plumbing is one that is widely debated and of practical importance. For example, biofilm is widely considered to be a key environment for proliferation of OPs (Temmerman et al., 2006), but release to bulk water

and conveyance via the tap is generally necessary to cause disease (Steinert et al., 2002). In this study, the high water use frequency was also characterized by the lowest proportion of shared OTUs and sequences between paired biofilm and bulk water sample (Fig. 3.2, Table B.4). This suggests that the shorter stagnation time translates into less time for microbial interaction between phases to occur and take root. Microbes released from the biofilm to the bulk water would also be flushed out more frequently at a high use tap. Interestingly, the effect of higher water use frequency was diminished by elevated water heater temperature ($T \geq 51°C$, Table B.5). Elevated temperature ($T \geq 51°C$) was associated with a slightly increased proportion of shared OTUs between the two phases (Table B.3), which could be related to an overall decrease in total OTUs at the tap (Fig. B.6). These findings have important implications for recommended sampling and monitoring protocols and whether biofilm or bulk water is most appropriately targeted over a range of conditions.

As noted above, the effect of gradually elevating the water heater temperature was not continuous, particularly in the bulk water. Notably, 51°C appeared to be a critical threshold, with phylogenetic dissimilarity among distal tap biofilm samples, alpha diversity in distal tap water samples, ammonia oxidation in recirculating line bulk water, and interactive effects with water use frequency all bifurcating around this temperature. One possibility is that 51°C demarcates a transition from a mesophilic to a thermophilic environment. For example, in anaerobic digestion, the mesophilic range is defined from 35-37°C, while thermophilic digestion occurs from 50-57°C, each imparting distinct dominant pathways for methanogenesis (Ali Shah et al., 2014). In sum, microbes must be biochemically adaptive, such that proteins, nucleic acids, and lipids can maintain function at hot temperatures. An earlier study of 204 complete genomes and proteomes reported increased purine load index (A+G content) in nucleotide composition as well as increased sum of fractions of IIe, Val, Tyr, Trp, Arg, Gly and Leu amino acids under elevated optimal growth temperatures (-10-110°C, Zeldovich et al., 2007).

We considered whether the “effective” temperature profile (i.e., the integral of instantaneous temperature over time between two sampling events) would provide a more meaningful predictor than the water heater temperature setting on observed phylogenetic dissimilarity patterns. Using this approach, distal tap sample compositions overlapped between those of the influent and

recirculating samples (Fig. B.4). Further, a negative association between "effective" temperature and microbial diversity (Shannon Index) was observed in distal tap and recirculating line bulk water samples. While we could find no prior domestic hot water system studies for comparison, our observation is consistent with previous surveys of geothermal systems as a natural analog. A negative linear relationship between three temperature groups (62-87°C) and the Shannon Index was reported for four sediment microbial communities in Great Boiling Spring (Cole et al., 2013). Similar negative association were documented along a temperature gradient of 52-75°C in a Japanese hot spring (Everroad et al., 2012). Importantly, in the present study, the "peak" alpha diversity occurred around the "effective" temperature of 20-25°C, which concurs with an extensive survey on 36 geothermal areas with 135 samples (Sharp et al., 2014), where they observed a bell-shape alpha diversity curve over a temperature range of 7.5-99°C with peak value at 24°C. Notably, the indoor built environment is commonly maintained from 20-25°C.

Given that it is impossible to eradicate microbes from water systems, a "probiotic", or more specifically a "prebiotic" framework has been proposed for manipulating the microbiome at the tap (Wang et al., 2013). Elevated temperature, in particular, can be highly effective for controlling OPs (Buse and Ashbolt, 2011; ANSI/ASHARE, 2015). For instance, *L. pneumophila* has an optimal growth range of 25-42°C (ASHRAE, 2000), with measurable deactivation at temperatures higher than 50°C (WHO, 2007). In fact, a recent study has proposed a diagnostic approach to inform potential *L. pneumophila* risk through closely monitoring temperature profiles at critical control points within the hot water distribution system (Bédard et al., 2015). On the other hand, elevated temperature is not universally effective, particularly when only temporarily applied. *Legionella* have been observed to proliferate via necrotrophic growth on heat-killed cells following thermal disinfection (Temmerman et al., 2006). Thus, the present study makes it clear that temperature settings have much more broad sweeping effects on the microbial composition at the tap, likely exerting indirect microbial ecological controls on *Legionella* proliferation. Here we observed that water heater temperature settings from 39-58°C were negatively associated with relative abundance of *Legionella*, but was less effective for *Legionella* control after cooling at the tap relative to the conditions of the consistently hot recirculating line. Further, at 51°C, markedly high relative abundance of *Legionella* OTUs were observed in the bulk water phase at low-use distal taps, a measurement consistent with high *L.*

*pneumophila* gene copy numbers determined by quantitative PCR (67 times higher on average than the inflow recirculating line water, Rhoads et al., 2015). It is speculated that, at 51°C, the heat shock served to diminish microbes competitive with *Legionella* (i.e., "probiotics"), creating an ecological "sweet spot" where *Legionella* could recover and proliferate with the extended stagnation (Rhoads et al., 2015). In fact, heat pre-treatment (50°C for 30 min) is widely applied to increase the specificity for isolating *Legionellae* from hot water systems (Leoni and Legnani, 2001).

Associations with *Legionella* spp. were examined to gain insight into potential antagonistic or synergistic relationships. In particular, a strong positive association was noted between the relative abundances of *Legionella* spp. and mitochondrial *V. vermiformis* (née *Hartmanella vermiformis*). Some free-living amoebae are either obligate or facultative pathogens (Brieland et al., 1997; Lorenzo-Morales et al., 2007) and can serve as vehicles/reservoirs for *L. pneumophila* (Wadowsky et al., 1988; Kuiper et al., 2004). Positive correlations between *V. vermiformis* and *Legionella* have been reported in treated and distributed waters (Valster et al., 2011; Wang et al., 2012), yet noted as conditional relationships and hypothesized to be dependent on strain type (Buse and Ashbolt, 2011) and in-building plumbing temperature (Rhoads et al., 2015). The relative abundance of mitochondrial *V. vermiformis* was also inversely associated with water heater temperature setting (39-58°C), where tolerance of this organism has been reported up to 53°C (Rohr et al., 1998). However, no spike in *Vermamoeba* relative abundance was observed in the 51°C low-use tap scenario, suggesting the role of other unknown microbial ecological relationships in influencing *Legionella*. Interestingly, *Mycobacterium* spp., a genus that contains some OP members, relative abundances were also found to be positively associated with those of *Legionella* spp. Previous studies have noted co-occurrence of *M. avium* complex and *Legionella* in water and/or biofilms in hot water systems ranging from 40-54°C (Bukh and Roslev, 2014) as well as in a drinking water distribution system (Whiley et al., 2014). Positive correlation between *Legionella* spp. and *Mycobacterium* spp. gene copies measured by quantitative PCR has also been reported in a drinking water distribution system (Lu et al., 2016).

The design of the rig employed in this study uniquely served to simulate building plumbing conditions in a controlled, replicated fashion and isolate the effect of individual variables. The

kind of insight gained would not be possible from a field study or simplistic bench-scale experiment alone, as the extreme complexities encountered from building to building make it impossible to identify causal factors. Nonetheless, it is important to note key differences in the design of the present study relative to the field. For instance, much longer distal pipes in the field would likely magnify impacts of pipe orientation and regrowth on levels of OPs in water. Real-world hot water systems will also have more complex water use patterns and employ a wide range of materials, which together will contribute to greater heterogeneity across distal tap microbiota. Overall this study demonstrates the possibility of implementing practical design and operation measures for intentionally controlling the microbiome composition at the tap, a topic of increasing interest not only for pathogen control, but also as linkages between the human microbiome and that of the built environment become more apparent.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENT

This work was supported by The Alfred P. Sloan MoBE program, the National Science Foundation (CBET Award #1336650) and the Institute for Critical Technology and Applied Science (ICTAS) at Virginia Tech. We thank Brian Badgley for lab supplies, and Kelly Peterson, Arjan Ahluwalia, Daniel Logsdon for their laboratory assistance.

## REFERENCES

Ali Shah F, Mahmood Q, Maroof Shah M, Pervez A, Ahmad Asad S. (2014). Microbial ecology of anaerobic digesters: the key players of anaerobiosis. *Sci World J* 204: 138752.

Allison SD, Martiny JBH. (2008). Resistance, resilience, and redundancy in microbial communities. *PNAS* 105 (suppl. 1): 11512 – 11519.

Anderson MJ. (2001). A new method for non-parametric multivariate analysis of variance. *Aust Ecol* 26: 32 – 46.

Anderson MJ. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* 62(1): 245 – 253.

ANSI/ASHRAE. (2015). Standard 188 Legionellosis: risk management for building water systems. Atlanta, Georgia, USA.

Ashbolt NJ. (2015). Microbial contamination of drinking water and human health from community water systems. *Curr Environ Health Rep* 2: 95 – 106.

ASHRAE. (2000). ASHRAE Guideline 12, Minimizing the Risk of Legionellosis Associated with Building Water Systems. Atlanta, Georgia, USA.

Bargellini A, Marchesi I, Righi E, Ferrari A, Cencetti S, Borella P, Rovesti S. (2011). Parameters predictive of *Legionella* contamination in hot water systems: association with trace elements and heterotrophic plate counts. *Water Res* 45(6): 2315 – 2321.

Baron JL, VikramA, Duda S, Stout JE, Bibby K. (2014). Shift in the microbial ecology of a hospital hot water system following the introduction of an on-site monochloramine disinfection system. *PLoS ONE* 9(7): e102679.

Battin TJ, Sloan WT, Kjelleberg S, Daims H, Head IM, Curtis TP, et al. (2007). Microbial landscapes: new paths to biofilm research. *Nat Rev Microbiol* 5: 7 – 12.

Bédard E, Fey S, Charron D, Lalancette C, Cantin P, Dolcé P, et al. (2015). Temperature diagnostic to identify high risk areas and optimize *Legionella pneumophila* surveillance in hot water distribution systems. *Water Res* 71: 244 – 256.

Beer KD, Gargano JW, Roberts VA, Hill VR, Garrison LE, Kutty PK, et al. (2015). Surveillance for waterborne disease outbreaks associated with drinking water – United States, 2011–2012. *Morb Mortal Wkly Rep* 64(31): 842 – 848.

Benjamini Y, Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B* 57(1): 289 – 300.

Borella P, Montagna MT, Romano-Spica V, Stampi S, Stancanelli G, Triassi M, et al. (2004). *Legionella* infection risk from domestic hot water. *Emerg Infect Dis* 10(3): 457 – 464.

Brieland JK, Fantone JC, Remick DG, LeGendre M, McClain M, Engleberg NC. (1997). The role of *Legionella pneumophila*-infected *Hartmannella vermiformis* as an infectious particle in a murine model of Legionnaires' disease. *Infect Immun* 65(12): 5330 – 5333.

Bukh AS, Roslev P. (2014). *Mycobacterium avium* complex in day care hot water systems, and persistence of live cells and DNA in hot water pipes. *Curr Microbiol* 68(4): 428 – 439.

Burrell PC, Phalen CM, Hovanec TA. (2001). Identification of bacteria responsible for ammonia oxidation in freshwater aquaria. *Appl Environ Microbiol* 67(12): 5791 – 5800.

Buse HY, Ashbolt NJ. (2011). Differential growth of *Legionella pneumophila* strains within a range of amoebae at various temperatures associated with in-premise plumbing. *Lett Appl Microbiol* 53(2): 217 – 224.

Cáceres MD, Legendre P. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology* 90(12): 3566 – 3574.

Cáceres MD, Legendre P, Moretti M. (2010). Improving indicator species analysis by combining groups of sites. *Oikos* 119(10): 1674 – 1684.

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

Chen H, Boutros PC. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* 12:35.

Cole JK, Peacock JP, Dodsworth JA, Williams AJ, Thompson DB, Dong H, et al. (2012). Sediment microbial communities in Great Boiling Spring are controlled by temperature and distinct from water communities. *ISME J* 7(4): 718 – 729.

Cooper IR, Hanlon GW. (2010). Resistance of *Legionella pneumophila* serotype 1 biofilms to chlorine-based disinfection. *J Hosp Infect* 74(2): 152 – 159.

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

Darelid J, Löfgren S, Malmvall BE. (2002). Control of nosocomial Legionnaires' disease by keeping the circulating hot water temperature above 55 degree C: experience from a 10-year surveillance programme in a district general hospital. *J Hosp Infect* 50(3): 213 – 219.

De Caceres M, Legendre P. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology* 90(12): 3566 – 3574.

Department of Veterans Affairs. (2014). Prevention of healthcare-associated *Legionella* disease and scald injury from potable water distribution systems. VHA Directive 1061.

Dumbrell AJ, Nelson M, Helgason T, Dytham C, Fitter AH. (2010). Relative roles of niche and neutral processes in structuring a soil microbial community. *ISME J* 4(3): 337 – 345.

El-Chakhtoura J, Prest E, Saikaly P, van Loosdrecht M, Hammes F, Vrouwenvelder H. (2015). Dynamics of bacterial communities before and after distribution in a full-scale drinking water network. *Water Res* 74(1): 180 – 190.

Everroad RC, Otaki H, Matsuura K, Haruta S. (2012). Diversification of bacterial community composition along a temperature gradient at a thermal spring. *Microbes Environ* 27(4): 374 – 381.

Falkinham III JO, Pruden A, Edwards MA. (2015). Opportunistic premise plumbing pathogens: increasingly important pathogens in drinking water. *Pathogens* 4(2): 373 – 386.

Farhat M, Moletta-Denat M, Frère J, Onillon S, Trouilhé M-C, Robine E. (2012). Effects of disinfection on *Legionella* spp., eukarya, and biofilms in a hot water system. *Appl Environ Microbiol* 78(19): 6850 – 6858.

Feazel LM, Baumgartner LK, Peterson KL, Frank DN, Harris JK, Pace NR. (2009). Opportunistic pathogens enriched in showerhead biofilms. *PNAS* 106(38): 16393 – 16399.

Fox J, Weisberg S. (2011). An R Companion to Applied Regression, Second Edition. SAGE Publications: Thousand Oaks, California, USA.

Hanski I, von Hertzen L, Fyhrquist N, Koskinen K, Torppa K, Laatikainen T, et al. (2012). Environmental biodiversity, human microbiota, and allergy are interrelated. *PNAS* 109(21): 8334 – 8339.

Hilborn ED, Wade TJ, Hicks L, Garrison L, Carpenter J, Adam E, et al. (2013). Surveillance for waterborne disease outbreaks associated with drinking water and other nonrecreational water—United States, 2009-2010. *Morb Mortal Wkly Rep* 62(35): 714 – 720.

Howe AD, Forster S, Morton S, Marshall R, Osborn KS, Wright P, et al. (2002). *Cryptosporidium* oocysts in a water supply associated with a cryptosporidiosis outbreak. *Emerg Infect Dis* 8(6): 619 – 624.

Hubbell SP. (2001). The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press: Princeton, New Jersey, USA.

Hutchinson GE. (1957). Concluding remarks. *Cold Spring Harbor Symp Quant Biol* 22: 415 – 427.

Jeraldo P, Sipos M, Chia N, Brulc JM, Dhillon AS, Konkel ME, et al. (2012). Quantification of the relative roles of niche and neutral processes in structuring gastrointestinal microbiomes. *PNAS* 109(25): 9692 – 9698.

Ji P, Parks J, Edwards MA, and Pruden A. (2015). Impact of water chemistry, pipe material and stagnation on the building plumbing microbiome. *PLoS ONE* 10(10): e0141087.

Kelley ST, Gilbert JA. (2013). Studying the microbiology of the indoor environment. *Genome Biol* 14(2): 202.

Kuiper MW, Wullings BA, Akkermans ADL, Beumer RR, van der Kooij D. (2004). Intracellular proliferation of *Legionella pneumophila* in *Hartmannella vermiformis* in aquatic biofilms grown on plasticized polyvinyl chloride. *Appl Environ Microbiol* 70(11): 6826 – 6833.

Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* 31: 814 – 821.

Lau HY, Ashbolt NJ. (2009). The role of biofilms and protozoa in *Legionella* pathogenesis: implications for drinking water. *J Appl Microbiol* 107(2): 368 – 378.

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

Lautenschlager K, Hwang C, Liu W-T, Boon N, Köster O, Vrouwenvelder H, et al. (2013). A microbiology-based multi-parametric approach towards assessing biological stability in drinking water distribution networks. *Water Res* 47(9): 3015 – 3025.

Leoni E, Legnani PP. (2001). Comparison of selective procedures for isolation and enumeration of *Legionella* species from hot water systems. *J Appl Microbiol* 90(1): 27 – 33.

Lévesque B, Lavoie M, Joly J. (2004). Residential water heater temperature: 49 or 60 degrees Celsius? *Can J Infect Dis* 15(1): 11 – 12.

Liu Z, Lin YE, Stout JE, Hwang CC, Vidic RD, Yu VL. (2006). Effect of flow regimes on the presence of *Legionella* within the biofilm of a model plumbing system. *J Appl Microbiol* 101(2): 437 – 442.

Lorenzo-Morales J, Martínez-Carretero E, Batista N, Alvarez-Marín J, Bahaya Y, Walochnik J, et al. (2007). Early diagnosis of amoebic keratitis due to a mixed infection with *Acanthamoeba* and *Hartmannella*. *Parasitol Res* 102(1): 167 – 169.

Lu J, Struewing I, Vereen E, Kirby AE, Levy K, Moe C, et al. (2016). Molecular detection of *Legionella* spp. and their associations with *Mycobacterium* spp., *Pseudomonas aeruginosa* and amoeba hosts in a drinking water distribution system. *J Appl Microbiol* 120(2): 509 – 521.

Lührig K, Canbäck B, Paul CJ, Johansson T, Persson KM, Rådström P. (2015). Bacterial community analysis of drinking water biofilms in southern Sweden. *Microbes Environ* 30(1): 99 – 107.

Martiny AC, Albrechtsen H-J, Arvin E, Molin S. (2005). Identification of bacteria in biofilm and bulk water samples from a non-chlorinated model drinking water distribution system: detection of a large nitrite-oxidizing population associated with *Nitrospira* spp. *Appl Environ Microbiol* 71(12): 8611 – 8617.

Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. (2010). PANDAseq: PAired-eND Assembler for Illumina sequences. *BMC Bioinformatics* 13: 31.

Mathys W, Stanke J, Harmuth M, Junge-Mathys E. (2008). Occurrence of *Legionella* in hot water systems of single-family residences in suburbs of two German cities with special reference to solar and district heating. *Int J Hyg Environ Health* 211(1-2): 179 – 185.

Oberauner L, Zachow C, Lackner S, Högenauer C, Smolle K-H, Berg G. (2013). The ignored diversity: complex bacterial communities in intensive care units revealed by 16S pyrosequencing. *Sci Rep* 3: 1413.

Ofiţeru ID, Lunn M, Curtis TP, Wells GF, Criddle CS, Francis CA, et al. (2010). Combined niche and neutral effects in a microbial wastewater treatment community. *PNAS* 107(35): 15345 – 15350.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. (2016). Vegan: Community Ecology Package. R package version 2.3-5. https://cran.r-project.org/web/packages/vegan/vegan.pdf.

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

Pohlert T. (2014). The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR). R package, http://CRAN.R-project.org/package=PMCMR.

Proctor CR, Gächter M, Kötzsch S, Rölli F, Sigrist R, Walser J-C, et al. (2016). Biofilms in shower hoses – choice of pipe material influences bacterial growth and communities. *Environ Sci: Water Res Technol* 2: 670 – 682.

Proctor CR, Hammes F. (2015). Drinking water microbiology – from measurement to management. *Curr Opin Biotechnol* 33: 87 – 94.

Pryor M, Springthorpe S, Riffard S, Brooks T, Huo Y, Davis G, et al. (2004). Investigation of opportunistic pathogens in municipal drinking water under different supply and treatment regimes. *Water Sci Technol* 50(1): 83 – 90.

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/, accessed on 02/29/2015.

Reshef DN, Reshef YA, Finucane HK, Grossman SR, McVean G, Turnbaugh PJ, et al. (2011). Detecting novel associations in large data sets. *Science* 334(6062): 1518 – 1524.

Rhoads WJ, Ji P, Pruden A, Edwards MA. (2015). Water heater temperature set point and water use patterns influence *Legionella pneumophila* and associated microorganisms at the tap. *Microbiome* 3: 67.

Rhoads WJ, Pruden A, Edwards MA. (2016a). Surveys of green building water systems reveals elevated water age and water quality concerns. *Environ Sci: Water Res Technol* 2: 164 – 173.

Rhoads WJ, Pruden A, Edwards MA. (2016b). Convective mixing in distal pipes exacerbates *Legionella pneumophila* growth in hot water plumbing. *Pathogens* 5(1): E29.

Rohr U, Weber S, Michel R, Selenka F, Wilhelm M. (1998). Comparison of free-living amoebae in hot water systems of hospitals with isolates from moist sanitary areas by identifying genera and determining temperature tolerance. *Appl Environ Microbiol* 64(5): 1822 – 1824.

Sharp CE, Brady AL, Sharp GH, Grasby SE, Stott MB, Dunfield PF. (2014). Humboldt's spa: microbial diversity is controlled by temperature in geothermal environments. *ISME J* 8: 1166 – 1174.

Steinert M, Hentschel U, Hacker J. (2002). *Legionella pneumophila*: an aquatic microbe goes astray. *FEMS Microbiol* Rev 26(2): 149 – 162.

Temmerman R, Vervaeren H, Noseda B, Boon N, Verstraete W. (2006). Necrotrophic growth of *Legionella pneumophila*. *Appl Environ Microbiol* 72(6): 4323 – 4328.

Valster RM, Wullings BA, van den Berg R, van der Kooij D. (2011). Relationships between free-living protozoa, cultivable *Legionella* spp., and water quality characteristics in three drinking water supplies in the Caribbean. *Appl Environ Microbiol* 77(20): 7321 – 7328.

Vázquez-Baeza Y, Pirrung M, Gonzalez A, Knight R. (2013). EMPeror: a tool for visualizing high-throughput microbial community data. *Gigascience* 2: 16.

Wadowsky RM, Butler LJ, Cook MK, Verma SM, Paul MA, Fields BS, et al. (1988). Growth-supporting activity for *Legionella pneumophila* in tap water cultures and implication of hartmannellid amoebae as growth factors. *Appl Environ Microbiol* 54(11): 2677 – 2682.

Wang H, Edwards MA, Falkinham III JO, 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. *Appl Environ Microbiol* 78(17): 6285 – 6294.

Wang H, Edwards MA, Falkinham III JO, Pruden A. (2013). Probiotic approach to pathogen control in premise plumbing systems? A review. *Environ Sci Technol* 47(18): 10117 – 10128.

Whiley H, Keegan A, Fallowfield H, Bentham R. (2014). Detection of *Legionella*, *L. pneumophila* and Mycobacterium Avium Complex (MAC) along potable water distribution pipelines. *Int J Environ Res Public Health* 11(7): 7393 – 7405.

WHO. (2007). *Legionella* and the prevention of legionellosis. http://www.who.int/water_sanitation_health/emerging/legionella_rel/en/. Accessed on July 1, 2016.

Zacheus OM, Martikainen PJ. (2004). Occurrence of legionellae in hot water distribution systems of Finnish apartment buildings. *Can J Microbiol* 40(12): 993 – 999.

Zeldovich KB, Berezovsky IN, Shakhnovich EI. (2007). Protein and DNA sequences determinants of thermophilic adaptation. *PLoS Comput Biol* 3(1): e5.

Zhang H-H, Chen S-N, Huang T-L, Shang P-L, Yang X, Ma W-X. (2015). Indoor heating drives water bacterial growth and community metabolic profile changes in building tap pipes during the winter season. *Int J Environ Res Public Health* 12(10): 13649 – 13661.

Zietz B, Wiese J, Brengelmann F, Dunkelberg H. (2001). Presence of *Legionellaceae* in warm water supplies and typing of strains by polymerase chain reaction. *Epidemiol Infect* 126(1): 147 – 152.

## APPENDIX B – SUPPLEMENTARY INFORMATION FOR CHAPTER 3

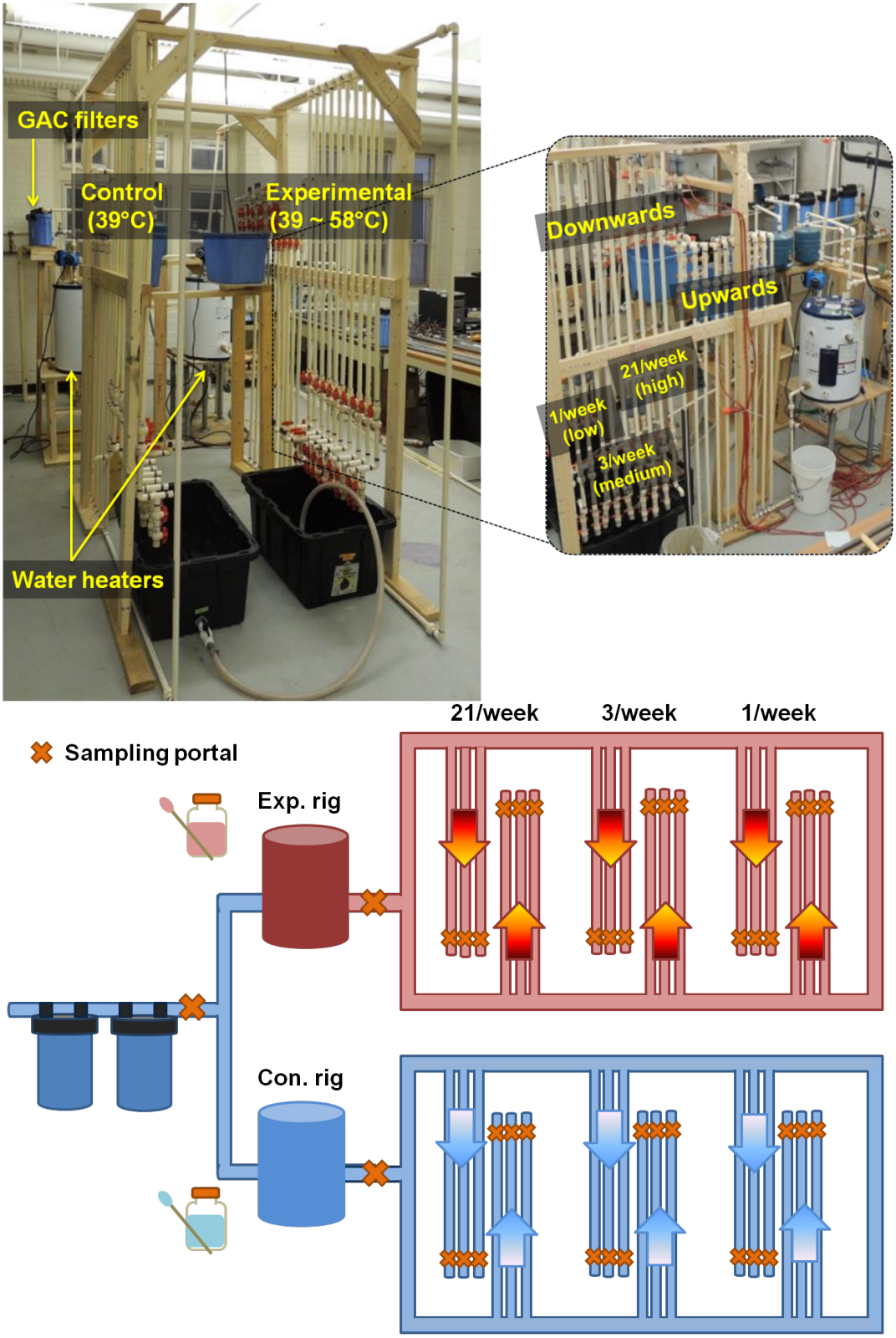


**Fig. B.1. Design of the lab-scale hot water heater rigs.**

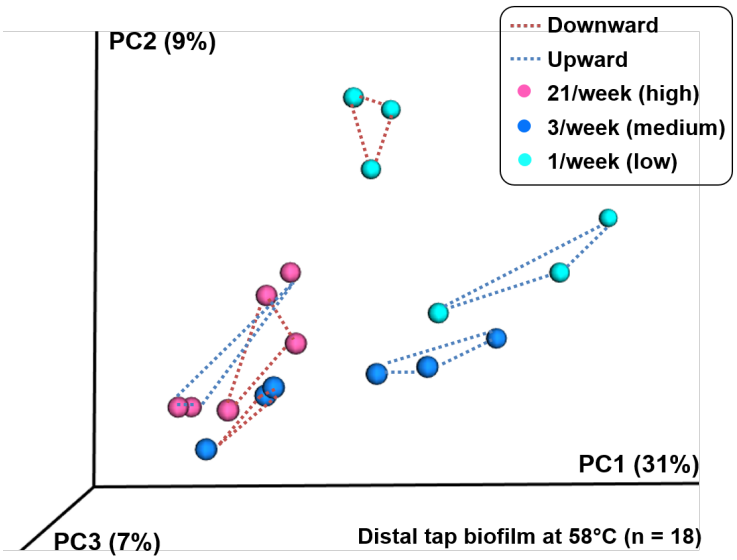


**Fig. B.2. Phylogenetic dissimilarity pattern among biofilm samples categorized by water use frequency** (colored points) **and pipe orientation** (dotted lines) (n=18, distal tap biofilm samples at water heater temperature setting of 58°C, 15 months)**.** Weighted UniFrac distance matrix was used (jackknifed 100 times).

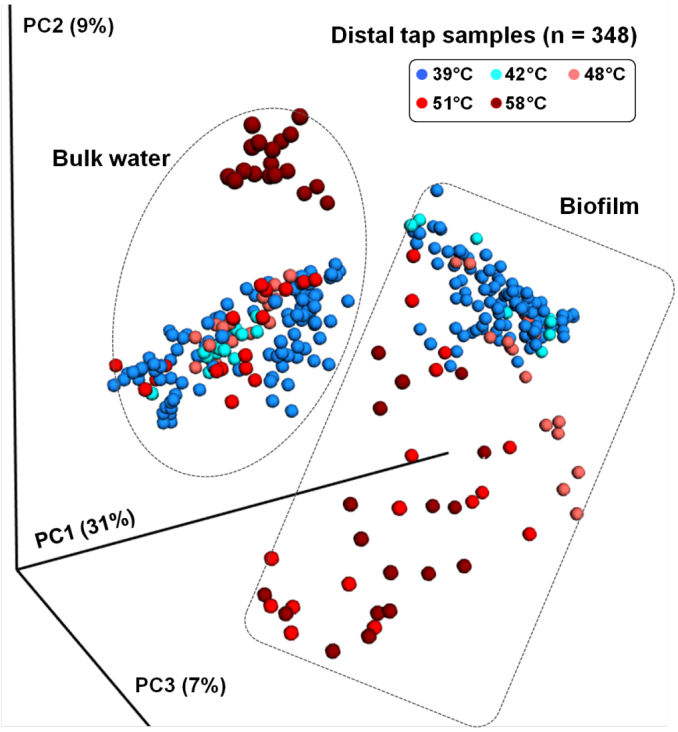


**Fig. B.3. Distal tap phylogenetic dissimilarity pattern across elevated water heater temperature setting (n = 348).** Weighted UniFrac distance matrix was used (jackknifed 100 times).

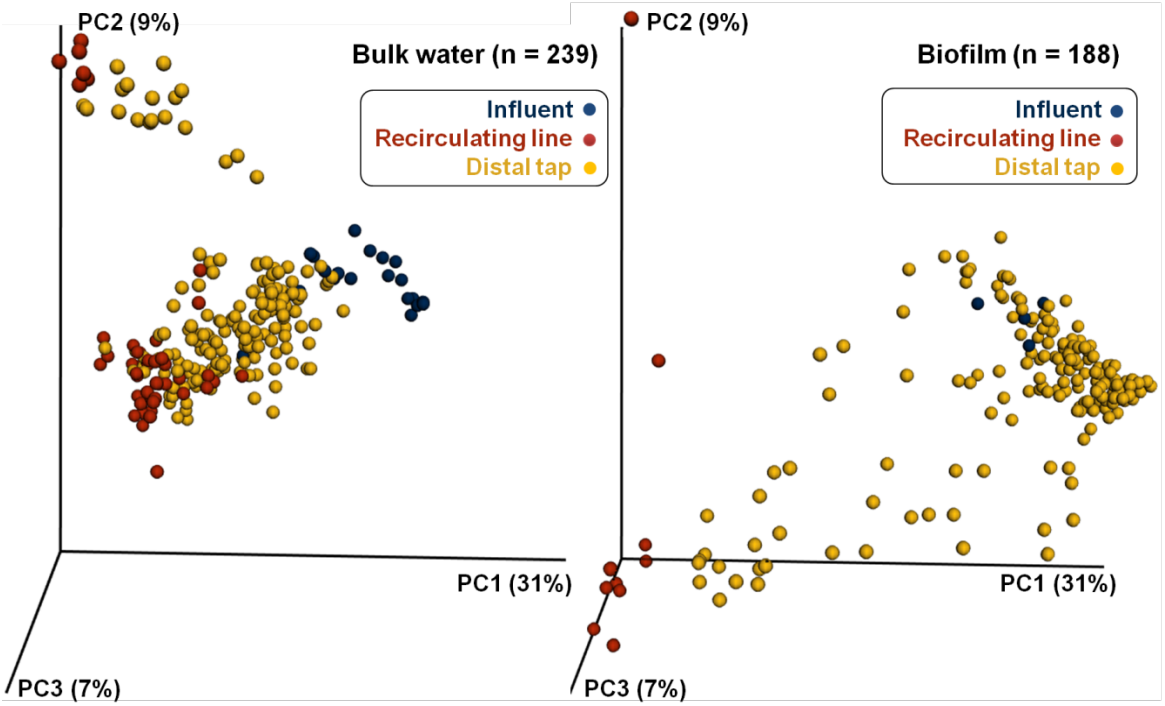


**Fig. B.4. 3D beta diversity plot of phylogenetic dissimilarity patterns comparing sampling location across all samples of each phase (bulk water n=239, biofilm n=188).**

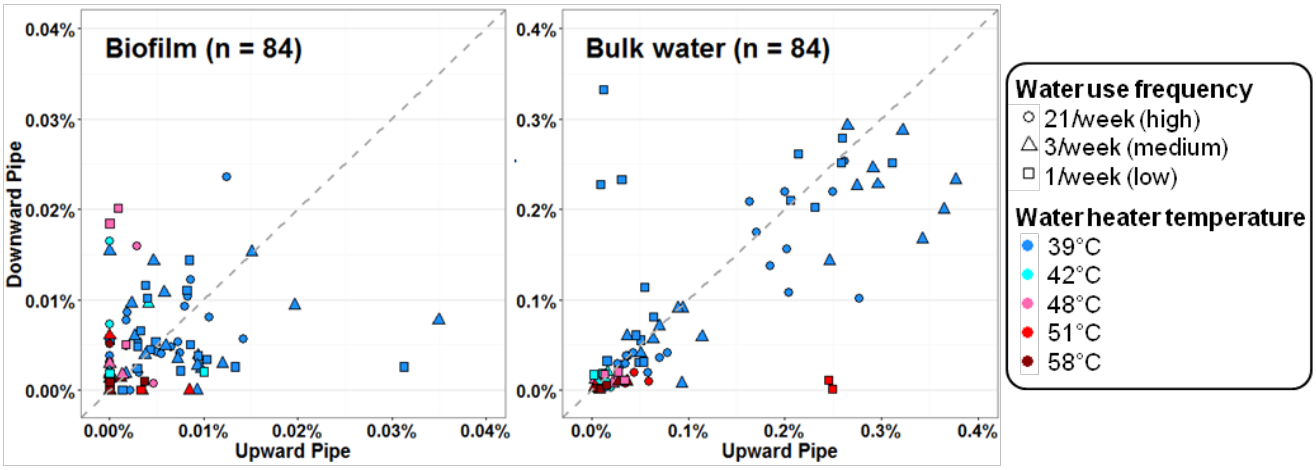


**Fig. B.5. Paired relative abundance of *Legionella* spp. in downward and upward pipes (n = 84 pairs of samples for each).** The grey dashed line demarcates a 1:1 ratio (downwards = upwards), indicating no effect of pipe orientation.

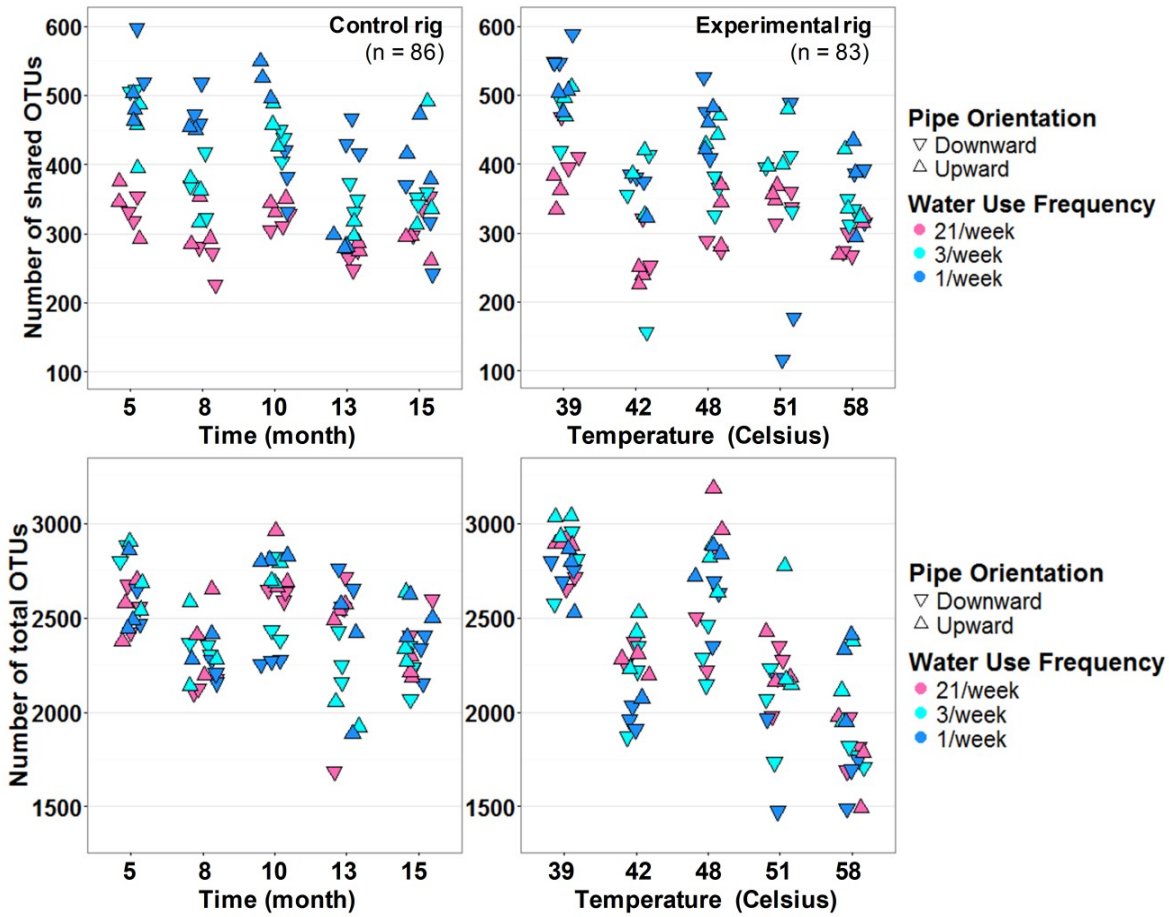


**Fig. B.6.** Number of shared and total OTUs in paired bulk water and biofilm samples at distal taps.

**Table B.1. Ammonia concentration across time and water heater temperature setting. Measurements were taken at influent and recirculating line sampling portals.** Net decrease in ammonia concentration was calculated as the difference between influent and corresponding recirculating line ones.

| Influent | Recirculating Line | | | | | |
|---|---|---|---|---|---|---|
| | Control rig | | | Experimental rig | | |
| Ammonia (mg/L) | t (months) | Ammonia (mg/L) | Net decrease (mg/L) | T (°C) | Ammonia (mg/L) | Net decrease (mg/L) |
| 0.56 | 5 | 0.57 | -0.01 | 39 | 0.49 | 0.07 |
| 0.48 | 8 | 0.16 | 0.32 | 42 | 0.15 | 0.33 |
| 0.41 | 10 | 0.08 | 0.33 | 48 | 0.10 | 0.31 |
| 0.42 | 13 | 0.06 | 0.36 | 51 | 0.34 | 0.08 |
| 0.46 | 15 | 0.08 | 0.38 | 58 | 0.36 | 0.10 |

**Table B.2. Relative abundance of *Nitrosomonas* spp. at recirculating line across time and water heater temperature setting.** Relative abundance was calculated using aggregated sequences of samples from the same group.

| Control rig | | | Experimental rig | | |
|---|---|---|---|---|---|
| t (months) | Bulk water | Biofilm | T (°C) | Bulk water | Biofilm |
| 5 | $3.88\times10^{-5}$ | 0 | 39 | $2.30\times10^{-5}$ | 0 |
| 8 | $7.79\times10^{-5}$ | 0 | 42 | $8.45\times10^{-6}$ | 0 |
| 10 | NA | 0 | 48 | NA | 0 |
| 13 | $2.16\times10^{-5}$ | 0 | 51 | $5.50\times10^{-6}$ | 0 |
| 15 | $3.51\times10^{-5}$ | 0 | 58 | $2.45\times10^{-6}$ | 0 |

**Table B.3. Impact of water heater temperature on the proportion of operational taxonomic units and sequences shared between biofilm and bulk water sample pairs.** Tukey Honestly Significant Differences test at 95% confidence intervals were applied to One-way analysis of variance results (n = 68 sample pairs at t = 13 months or 15 months in both rigs). Statistical significance was set at 0.05 and italic grey font suggested non-statistically significant.

| Groups | Proportion of shared OTUs | | Proportion of shared sequences | |
|---|---|---|---|---|
| | $\Delta_{Mean}$ | $P_{adjusted}$ | $\Delta_{Mean}$ | $P_{adjusted}$ |
| $51°C_{13\ month}$ vs $39°C_{13\ month}$ | 0.026 | 0.046 | *0.004* | *0.992* |
| $58°C_{15\ month}$ vs $39°C_{15\ month}$ | 0.031 | 0.005 | *0.026* | *0.258* |
| $39°C_{15\ month}$ vs $39°C_{13\ month}$ | *0.010* | *0.717* | *0.018* | *0.559* |
| $58°C_{15\ month}$ vs $39°C_{13\ month}$ | 0.041 | <0.0002 | 0.044 | 0.013 |
| $51°C_{13\ month}$ vs $39°C_{15\ month}$ | *0.016* | *0.341* | *-0.014* | *0.759* |
| $58°C_{15\ month}$ vs $51°C_{13\ month}$ | *0.016* | *0.357* | *0.040* | *0.038* |
| Levene's test Pr (>F) | 0.392 | | 0.062 | |

**Table B.4. Impact of flushing frequency on the proportion of operational taxonomic units and sequences shared between biofilm and bulk water sample pairs** (n = 169)**.** Kruskal and Wallis test with Tukey and Kramer (Nemenyi) test was applied with statistical significance set at 0.05 and non-statistically significant ones in italic grey font.

| Groups | Proportion of shared OTUs | | Proportion of shared sequences | |
|---|---|---|---|---|
| | $\Delta_{Median}$ | $P_{adjusted}$ | $\Delta_{Median}$ | $P_{adjusted}$ |
| 3/week (high) - 21/week (medium) | 0.03378 | <0.0001 | 0.058321 | <0.0001 |
| 1/week (high) - 21/week (low) | 0.05115 | <0.0001 | 0.062514 | <0.0001 |
| 1/week (medium) - 3/week (low) | *0.01738* | *0.14* | *0.004193* | *0.97* |
| Levene's test Pr (>F) | < 0.001 | | < 0.0001 | |

**Table B.5. Impact of flushing frequency on the proportion of operational taxonomic units and sequences shared between biofilm and bulk water sample pair from each rig and time/temperature.** One-way ANOVA with Tukey HSD test or Kruskal and Wallis test with Tukey and Kramer (Nemenyi) test was applied depending on Levene's test results. Statistical significance set at 0.05, with non-statistically significant values indicated in italic grey font.

| **Proportion of shared OTUs (control rig)** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Water use frequency** | **5 months** | | **8 months** | | **10 months** | | **13 months** | | **15 months** | |
| | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** |
| 3/week(medium) - 21/week (high) | 0.039 | 0.001 | 0.030 | 0.003 | 0.047 | <0.001 | 0.040 | 0.007 | *0.025* | *0.059* |
| 1/week (low) - 21/week (high) | 0.067 | <0.001 | 0.083 | <0.001 | 0.055 | <0.001 | 0.031 | 0.024 | *0.019* | *0.173* |
| 1/week (low) - 3/week (medium) | 0.029 | 0.016 | 0.054 | <0.001 | *0.008* | *0.570* | *-0.008* | *0.735* | *-0.006* | *0.818* |
| Levene's test Pr (>F) | *0.444* | | *0.613* | | *0.178* | | *0.150* | | *0.202* | |
| Sample size (pairs) | 16 | | 17 | | 18 | | 17 | | 18 | |
| **Proportion of shared OTUs (experimental rig)** | | | | | | | | | | |
| **Water use frequency** | **39°C** | | **42°C** | | **48°C** | | **51°C** | | **58°C** | |
| | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Median}$** | **$P_{adjusted}$** |
| 3/week(medium) - 21/week (high) | 0.031 | 0.009 | *0.036* | *0.083* | 0.045 | <0.001 | *0.029* | *0.341* | *0.015* | *0.284* |
| 1/week (low) - 21/week (high) | 0.054 | <0.001 | 0.071 | 0.003 | 0.059 | <0.001 | *-0.025* | *0.562* | *0.031* | *0.099* |
| 1/week (low) - 3/week (medium) | 0.024 | 0.040 | *0.036* | *0.111* | *0.014* | *0.120* | *-0.054* | *0.103* | *0.016* | *0.851* |
| Levene's test Pr (>F) | *0.587* | | *0.702* | | *0.241* | | *0.161* | | 0.009 | |
| Sample size (pairs) | 18 | | 15 | | 17 | | 15 | | 18 | |
| **Proportion of shared sequences (control rig)** | | | | | | | | | | |
| **Water use frequency** | **5 months** | | **8 months** | | **10 months** | | **13 months** | | **15 months** | |
| | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Median}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** |
| 3/week(medium) - 21/week (high) | 0.085 | <0.001 | 0.084 | <0.001 | 0.095 | 0.034 | 0.067 | 0.035 | 0.034 | 0.003 |
| 1/week (low) - 21/week (high) | 0.107 | <0.001 | 0.1012 | <0.001 | 0.098 | 0.002 | *0.029* | *0.430* | 0.028 | 0.010 |
| 1/week (low) - 3/week (medium) | *0.022* | *0.43* | *0.018* | *0.441* | *0.004* | *0.662* | *-0.038* | *0.284* | *-0.005* | *0.790* |
| Levene's test Pr (>F) | *0.288* | | *0.063* | | 0.001 | | *0.173* | | *0.573* | |

| Sample size (pairs) | 16 | | 17 | | 18 | | 17 | | 18 | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Proportion of shared sequences (experimental rig)** | | | | | | | | | | |
| **Water use frequency** | **39°C** | | **42°C** | | **48°C** | | **51°C** | | **58°C** | |
| | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Median}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** | **$\Delta_{Mean}$** | **$P_{adjusted}$** |
| 3/week(medium) - 21/week (high) | 0.048 | 0.002 | *0.051* | *0.110* | 0.093 | 0.014 | *0.038* | *0.463* | *0.007* | *0.822* |
| 1/week (low) - 21/week (high) | 0.067 | <0.001 | 0.078 | 0.023 | *0.093* | *0.071* | *-0.080* | *0.136* | *0.006* | *0.862* |
| 1/week (low) - 3/week (medium) | *0.018* | *0.266* | *0.027* | *0.528* | *0.0003* | *0.804* | -0.118 | 0.024 | *-0.0009* | *0.997* |
| Levene's test Pr (>F) | *0.143* | | *0.728* | | 0.049 | | *0.051* | | *0.069* | |
| Sample size (pairs) | 18 | | 15 | | 17 | | 15 | | 18 | |

# CHAPTER 4. EFFECT OF HEAT SHOCK ON HOT WATER PLUMBING MICROBIOTA AND *LEGIONELLA PNEUMOPHILA* CONTROL

Pan Ji, William J. Rhoads, Marc A. Edwards, Amy Pruden

## ABSTRACT

**Background**: Heat shock is a potential control strategy for *Legionella pneumophila* in hot water plumbing systems. However, it is not consistently effective, with little understanding of its influence on the broader plumbing microbiome. Here we employed a lab-scale recirculating hot water plumbing rig to compare the pre- and post- "heat shock" (i.e., 40 → 60 → 40 °C) microbiota at distal taps. In addition, we used a second plumbing rig to represent a well-managed system at 60 °C and conducted a "control" sampling at 60 °C, subsequently reducing the temperature to 40 °C to observe the effects on *Legionella* and the microbiota under a simulated "thermal disruption" scenario.

**Results:** According to 16S rRNA gene amplicon sequencing, in the heat shock scenario, there was no significant difference or statistically significant, but small, difference in the microbial community composition at the distal taps pre- versus post- heat shock (both biofilm and water; weighted and unweighted UniFrac distance matrices). While heat shock did lead to decreased total bacteria numbers at distal taps, it did not measurably alter the richness or evenness of the microbiota. Quantitative PCR measurements demonstrated that *L. pneumophila* relative abundance at distal taps also was not significantly different at 2-month post- heat shock relative to the pre- heat shock condition, while relative abundance of *Vermamoeba vermiformis*, a known *Legionella* host, did increase. In the thermal disruption scenario, relative abundance of planktonic *L. pneumophila* (quantitative PCR data) increased to levels comparable to those observed in the heat shock scenario within 2 months of switching long-term operation at 60 to 40 °C. Overall, water use frequency and water heater temperature set point exhibited a stronger effect than one-time heat shock on the microbial composition and *Legionella* levels at distal taps.

**Conclusions:** While heat shock may be effective for instantaneous *Legionella* control and reduction in total bacteria numbers, water heater temperature set point and water use frequency are more promising factors for long-term *Legionella* and microbial community control, illustrating the importance of maintaining consistent elevated temperatures in the system relative to short-term heat shock.

**KEYWORDS**

Heat shock - Hot water plumbing - Distal taps - Opportunistic pathogens – Biofilm - 16S rRNA gene amplicon sequencing

## INTRODUCTION

Hot water systems are a key source of microbes to the human-occupied built environment and harbor distinct microbiota from that of influent potable water (Ji et al., 2017). In particular, while influent cold water lines can be subject to significant seasonal variation (Pinto et al., 2014), spatial and temporal patterns in the microbial community composition of hot water systems can be even more complex due to variable flow patterns and configurations (Douterelo et al., 2013), stagnation time of distal taps (Lautenschlager et al., 2010; Ji et al., 2015), and temperature conditions (Ji et al., 2017). Further, in-building hot water systems often involve storage equipment (e.g., water heater or hot water tank), which can serve as a reservoir for microbes and contribute to downstream warm water conditions conducive to microbial growth (Rhoads et al., 2016). Elevated temperatures also accelerate disinfectant decay (e.g., chlorine, Hua et al., 1999) and predispose hot water systems to deteriorating microbial water quality. Such unique aspects of hot water systems, together with the inherent heterogeneity in domestic plumbing designs across different buildings, make it almost impossible to promote a unified control strategy for microbial regrowth.

Hot water systems are especially vulnerable to the growth of opportunistic pathogens (OPs), such as *Legionella pneumophila* and *Mycobacteria avium.* This emphasizes the critical role of design and operation for protecting public health, particularly when serving immunocompromised populations, such as in hospital settings. As early as 1987, hospital hot water tanks with temperature settings below 55 °C were identified as the primary source for nosocomial Legionnaire's disease outbreaks (Ware, 1989; Darelid et al., 2002). Notably, implementation of energy and water conservation features can unintentionally increase risk of OPs exposure (Nguyen et al., 2012), as was observed in a 400-bed university hospital in Sherbrooke, Canada, where elevated *L. pneumophila* growth in the hot water system was associated with the installation of a heat exchanger (Bédard et al., 2016b). In residential homes, showers represent a routine source of potential exposure to aerosolized OPs. Importantly, inhalation of such aerosols is the primary route of infection, rather than ingestion, as is the emphasis of drinking water regulations (EPA, 2016).

Much attention has been directed to control measures for OPs in hot water systems, including thermal disinfection/ heat shock (Bédard et al., 2016a), UV disinfection (Pozos et al., 2004), on-site secondary disinfection (e.g., monochloramine, Baron et al., 2015), and copper-silver ionization (Stout and Yu, 2003). Still, there is wide debate regarding the optimal choice for on-site OP control, with thermal disinfection or heat shock remaining one of the most widely accessible and feasible options for many building owners and residents. Thermal disinfection, or heat shock, typically involves setting the water heater temperature at a high level over a defined period of time and subsequently continuously or periodically flushing distal taps for a target duration at a minimal at-the-tap flushing temperature. Notably, there is a range of thermal disinfection or heat shock procedures defined by various professional and public health agencies with respect to several key elements (see Table C.1), including water heater temperature set point (60 – 77 °C) and flushing conditions (continuous or periodic).

Control of OPs within hot water systems is inherently a matter of managing the microbial ecology (Wang et al., 2013), as OPs are native to the drinking water environment and thus not as readily eradicated as fecal pathogens. Further, the high surface area to volume ratio characteristic of domestic plumbing encourages biofilm development, where OPs, including *L. pneumophila*, benefit from a parasitic relationship with amoebae that enables their proliferation within the highly oligotrophic drinking water environment (Declerck, 2010; Falkinham et al., 2015). However, the precise effects of heat shock for control of OPs (especially *L. pneumophila*) have not been systematically evaluated. Prior studies have examined *L. pneumophila* specifically (Yu et al., 1982; Bédard et al., 2016a) and short-term response to heat shock (e.g., 7 days, Farhat et al., 2012) or heat-treated tap water (Vervaeren et al., 2006). Still, significant knowledge gaps remain with respect to long-term effects on: a) OPs within the context of the broader microbial community composition; and b) bulk water and biofilm phases and their interrelationship.

This study employed a heat shock protocol at the "mild" end of the spectrum, with water heater temperature set point elevated to 60 °C and periodic flushing at distal taps to maintain at-the-tap temperature > 55 °C for 30 minutes, to gain a sense of the physical effect of heat shock to the microbes within a temperature regime widely accessible to building owners and residents. Here

we employed a lab-scale recirculating hot water plumbing rig to compare the pre- and post-heat shock (i.e., 40 → 60 → 40 °C) microbiota at distal taps ("heat-shock" scenario). In addition, we used a second plumbing rig to represent a well-managed system maintaining elevated temperature throughout the recirculating line (60 °C), and reduced the temperature to 40 °C to observe the effects on *Legionella* and the microbiota ("thermal disruption" scenario). Effects of heat shock and thermal disruption were compared relative to those imparted by the water heater temperature set point, pipe orientation and the water use frequency at the tap.

## METHODS

**System setup and experimental design.** Two identical hot water rigs were constructed to examine the impact of thermal conditions. Rig design has been previously described in detail (Rhoads et al., 2016; Rhoads et al., 2017). Each rig consisted of an electric water heater and recirculating pipe, with 18 distal taps comparing two pipe orientations (downward with little convective-mixing vs slanted upward with enhanced convective-mixing) and three water use frequencies in triplicate (high-, medium-, low- water use as 21, 3, 1 flushes/week respectively). Prior to this study, both rigs had been acclimated with Blacksburg, VA tap water for 15 months. Municipal chloramine residual was removed by passing water through three granular activated carbon filters in series (Rhoads et al., 2015; Ji et al., 2017). In addition, the upward oriented pipes were tilted 30° from vertical 4-month pre- heat shock (2 months prior to the first sampling point) to induce convective mixing for comparison to downward oriented pipes without convective mixing.

**Heat shock and thermal disruption.** Parallel comparison of two rigs allowed examination of heat shock as a control measure relative to water heater temperature set point and flow conditions at the tap. Prior to imposing the shifts in thermal conditions associated with the present study, the "heat shock" (referred to as "control" in prior studies) and "thermal disruption" (referred to as "experimental" in prior studies) rigs had been maintained at water heater temperature set points of 40 and 60 °C, respectively, for 4 months (Ji et al., 2017; Rhoads et al., 2016; Rhoads et al., 2017). Note that all temperatures cited herein can vary from the water heater temperature set point by ±1-2 °C. To commence the present study, both rigs were set to

60 °C and each set of distal taps were flushed intermittently to maintain water temperatures > 55 °C for approximately 30 minutes, targeting the guidance of the Association of Water Technologies (Association of Water Technologies (AWT), 2003) and Stout et al. (1986) (summary of published heat shock treatment procedures summarized in Table C.1, adapted from Table 1 in Rhoads et al., 2014). While this did not constitute a substantive 'heat shock' to the "thermal disruption" rig (given that it was already maintained at 60 °C), both rigs were subject to the same treatment to normalize the effects of flushing the distal lines at elevated temperature for an extended period of time. Post- heat shock, both water heater temperatures were set to 40 °C.

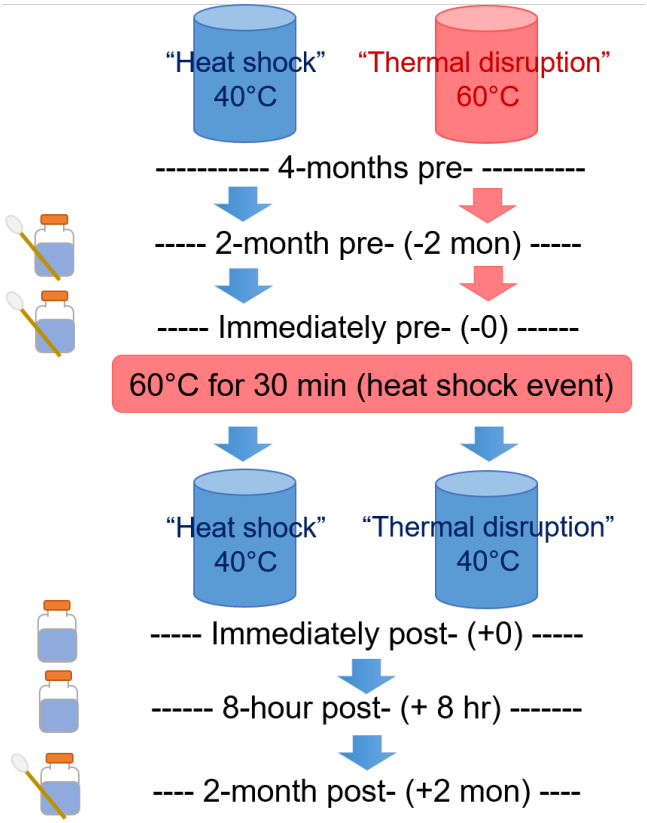


**Fig. 4.1. Experimental timeline.** Time is relative to the heat shock event.

**Sample collection.** First-flush bulk water (~500 mL) and biofilm (65 $cm^2$ inner surface area swabbed) samples were collected at each sampling portal (influent, recirculating line and distal taps) 2-month pre- (i.e., time = -2 mon), immediately pre- (time = 0) and 2-month post- heat shock (time = +2 mon) (Fig. 4.1). Additional first-flush bulk water samples were collected immediately post- (+0) and 8-hour post- (+8 hr) heat shock (Fig. 4.1). The rationale for

collecting these additional bulk water samples was to capture potential changes during the heat shock process (immediately post heat shock) and after a modest stagnation time mimicking daily water use pattern (8-hour post- heat shock). If not specified, the term 'pre-heat shock samples' refers to samples taken at 2-month pre- and immediately pre- heat shock, while the term 'post-heat shock samples' only includes samples taken at 2-month post- heat shock. Distal tap samples were typically collected at the end of the cyclical 8-hour stagnation periods for each water use frequency.

**DNA extraction, qPCR, and 16S rRNA amplicon sequencing.** ***DNA extraction.*** Bulk water samples were first concentrated onto sterile 0.22-μm pore size mixed-cellulose-ester filters (Millipore, Billerica, MA, USA). DNA extraction from the fragmented filters and cotton swabs followed the FastDNA Spin Kit (MP Biomedicals, Solon, OH, USA) manufacturer protocol. ***Quantitative PCR.*** Quantitative Polymerase Chain Reaction (qPCR) was applied to quantify the gene copy numbers of total bacteria (16S rRNA), *L. pneumophila*, *Mycobacterium avium* and *Vermamoeba vermiformis* (née *Hartmanella vermiformis*) using established protocols (Wilton and Cousins, 1992; Suzuki et al., 2000; Kuiper et al., 2006; Nazarian et al., 2008; Wang et al., 2012). A dilution ratio of 1:10 was selected for all DNA extracts to balance inhibition and detection. Each sample was analyzed in triplicate, where at least two positive reads were scored as a positive detection of a given gene. To determine relative abundances, *L. pneumophila*, *M. avium* and *V. vermiformis* gene copy numbers were normalized to total bacterial 16S rRNA gene copy numbers. ***Amplicon sequencing.*** Sample preparation for 16S rRNA gene amplicon sequencing followed the online Earth Microbiome Project protocol (Earth Microbiome Project: http://www.earthmicrobiome.org/, accessed April 02, 2017) using the 515F/926R primer pair targeting the V4 and V5 regions of the 16S rRNA gene. Minor differences include using molecular grade water (Quality Biological, Gaithersburg, MD, USA) and pooling PCR products on an equal mass basis of 200 ng. Illumina amplicon sequencing was performed on MiSeq platform at the Biocomplexity Institute at Virginia Tech (paired-end 300 bp reads using MiSeq Kit V3).

**Amplicon sequencing data analysis.** Demultiplexed amplicon sequencing data were retrieved and processed using the PANDAseq assembler (Masella et al., 2012) to stitch the paired-end

reads with the criteria that the stitched read length should be between 372-375 bp and the threshold score of at least 0.80. Chimera-free sequences (USEARCH v6.1 (Edgar, 2010), reference-based chimera detection using Greengene database v13_8 (McDonald et al., 2012)) were subject to *de novo* operational taxonomy unit (OTU) picking strategy (pick_de_novo_otus.py) at 0.97 similarity (UCLUST, Edgar, 2010) in QIIME 1.8.0 (Caporaso et al., 2010b) referencing Greengene database v13_8 (McDonald et al., 2012). Sequences were aligned using PyNAST (Caporaso et al., 2010a). Taxonomy was assigned using RDP Classifier 2.2 (Wang et al., 2007). An approximately-maximum-likelihood phylogenetic tree was constructed using FastTree 2.1.3 (Price et al., 2010). Further, singletons (OTU with 1 sequence across the entire OTU table) and organelle OTUs (chloroplast and mitochondria) were removed from downstream analysis. A total of 10, 313, 752 sequences were retained for all 323 samples with a median value of 31, 946 sequences per sample (min: 5, 210, max: 111, 018). The cleaned OTU table was then rarefied 100 times to a sequencing depth of 5, 200, from which alpha diversity (Chao 1 index for richness; Gini index for evenness) and beta diversity (weighted and unweighted UniFrac distance matrices, Lozupone and Knight, 2005) was measured. Difference in alpha diversity (Chao 1 index) across different time points pre- and post- heat shock was examined via Kruskal-Wallis test, with Nemenyi test for pairwise comparisons (package "PMCMR" version 4.1, Pohlert, 2016). Permutational Multivariate Analysis of Variance (Adonis, Anderson, 2001) was applied to the average weighted and unweighted UniFrac distance matrices as a measurement of difference in group means. Complementary multivariate homogeneity of group dispersions analysis (betadisper, Anderson, 2006) was applied to evaluate the within-group variations (package "vegan" version 2.3-0, Oksanen et al., 2016), where homogeneity of dispersion among groups is an assumption for Adonis. All sequence data have been deposited in QIITA under Study ID 10504 and European Nucleotide Archive (ENA) under accession number PRJEB22241.

**Q-PCR data analysis.** Gene copy numbers determined by qPCR were converted to concentration in water/ biofilm based on the volume/ area sampled. Gene copy numbers were normalized with total bacterial 16S rRNA gene copy numbers as a proxy indicator of relative abundance. All statistical analysis, including linear regression and Spearman Correlation Analysis was conducted in R (version 3.3.1, R Core Team, 2016).

## RESULTS

**Heat shock imposed limited impact on microbial composition of biofilm.** For the biofilm phase, weighted and unweighted UniFrac distance matrices yielded highly similar trends (Table 4.1a, Table 4.1b; Fig. C.1). In the "heat shock" rig, comparison of pre- vs post- heat shock biofilm samples yielded either statistically significant, but minute, difference (2-month pre- vs 2-month post-, weighted UniFrac, $R^2_{Adonis}$ = 0.0760, $P_{Adonis}$ = 0.020) or no statistically significant difference (immediately pre- vs 2-month post-, weighted UniFrac, $R^2_{Adonis}$ = 0.0623, $P_{Adonis}$ = 0.056; Table 4.1a) in microbial community composition. In the "thermal disruption" rig, post-heat shock (40 °C) biofilm samples appeared to be distinct from those pre- heat shock (60 °C) (weighted UniFrac, $R^2_{Adonis}$ = 0.2104/0.2193, $P_{Adonis}$ = 0.001, Table 1b), where temporal variation within pre- heat shock biofilm samples were insignificant (weighted UniFrac, $R^2_{Adonis}$ = 0.0202, $P_{Adonis}$ = 0.566; Table 1b). Together these results are suggestive that the water heater set point had a stronger influence than the prior heat shock history (Table 4.1a, Table 4.1b).

**Table 4.1. Statistical comparison of samples collected pre- and post- heat shock (permutation = 999).**

| **a. "Heat shock" rig, biofilm samples** | **Weighted UniFrac** | | | **Unweighted UniFrac** | | |
|---|---|---|---|---|---|---|
| | $R^2_{Adonis}$ | $P_{Adonis}$ | $P_{Betadisp}$ | $R^2_{Adonis}$ | $P_{Adonis}$ | $P_{Betadisp}$ |
| 2-month pre- (19) vs 2-month post- (19) | 0.0760 | 0.020 | 0.919 | 0.0737 | 0.001 | 0.430 |
| 2-month pre- (19) vs immediately pre- (18) | 0.0310 | 0.307 | 0.257 | 0.0352 | 0.098 | 0.901 |
| Immediately pre- (18) vs 2- month post- (19) | 0.0623 | 0.056 | 0.297 | 0.0553 | 0.001 | 0.416 |
| **b. "Thermal disruption" rig, biofilm samples** | **Weighted UniFrac** | | | **Unweighted UniFrac** | | |
| | $R^2_{Adonis}$ | $P_{Adonis}$ | $P_{Betadisp}$ | $R^2_{Adonis}$ | $P_{Adonis}$ | $P_{Betadisp}$ |
| 2-month pre- (19) vs 2-month post- (19) | 0.2104 | 0.001 | 0.169 | 0.1525 | 0.001 | 0.001 |
| 2-month pre- (19) vs immediately pre- (18) | 0.0202 | 0.566 | 0.781 | 0.0292 | 0.299 | 0.995 |
| Immediately pre- (18) vs 2-month post- (19) | 0.2193 | 0.001 | 0.128 | 0.1504 | 0.001 | 0.001 |
| **c. "Heat shock" rig, water samples** | **Weighted UniFrac** | | | **Unweighted UniFrac** | | |
| | $R^2_{Adonis}$ | $P_{Adonis}$ | $P_{Betadisp}$ | $R^2_{Adonis}$ | $P_{Adonis}$ | $P_{Betadisp}$ |
| 2-month pre- (20) vs 2-month post- (21) | 0.1434 | 0.003 | 0.018 | 0.0789 | 0.001 | 0.019 |
| 2-month pre- (20) vs immediately pre- (18) | 0.2274 | 0.001 | 0.400 | 0.0598 | 0.001 | 0.604 |
| Immediately pre- (18) vs 2-month post- (21) | 0.2209 | 0.001 | 0.006 | 0.0763 | 0.001 | 0.005 |
| **d. "Thermal disruption" rig, water samples** | **Weighted UniFrac** | | | **Unweighted UniFrac** | | |
| | $R^2_{Adonis}$ | $P_{Adonis}$ | $P_{Betadisp}$ | $R^2_{Adonis}$ | $P_{Adonis}$ | $P_{Betadisp}$ |
| 2-month pre- (20) vs 2-month post- (21) | 0.3808 | 0.001 | 0.340 | 0.1737 | 0.001 | 0.499 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2-month pre- (20) vs immediately pre- (18) | 0.1038 | 0.009 | 0.471 | 0.0455 | 0.003 | 0.478 |
| Immediately pre- (18) vs 2-month post- (21) | 0.2498 | 0.001 | 0.056 | 0.1593 | 0.001 | 0.979 |

**Heat shock imposed limited impact on microbial composition of bulk water.** For the bulk water phase, the impacts of the experimental conditions were generally more apparent when applying the weighted (Fig. C.1), than the unweighted, UniFrac distance matrix (Table 4.1c, 4.1d), indicating a stronger effect in terms of relative abundance than occurrence of OTUs. In the "heat shock" rig, although the 2-month post-heat shock microbial community structure was significantly different from that of the pre- heat shock condition (weighted UniFrac, $R^2_{Adonis}$ = 0.1434/ 0.2209, $P_{Adonis}$ = 0.003/ 0.001), the difference observed was no greater than that associated with temporal variation of the two pre- heat shock samplings (2-month pre- vs immediately pre-, weighted UniFrac, $R^2_{Adonis}$ = 0.2274, $P_{Adonis}$ = 0.001, Table 4.1c). In the "thermal disruption" rig, there was a much sharper difference between the pre- and post-temperature drop water samples (weighted UniFrac, $R^2_{Adonis}$ = 0.3808/ 0.2498, $P_{Adonis}$ = 0.001) than between the two water samples collected at the 60°C set point before the temperature drop to 40°C (weighted UniFrac, $R^2_{Adonis}$ = 0.1038, $P_{Adonis}$ = 0.009). This again points to the water heater set point as the dominant factor shaping the microbial community structure.

Eight-hour stagnation (immediately post- vs 8-hr post-) was found to incur a statistically significant, but small, change in microbial composition based on both weighted ($R^2_{Adonis}$ = 0.0774, $P_{Adonis}$ = 0.001) and unweighted ($R^2_{Adonis}$ = 0.0371, $P_{Adonis}$ = 0.001) UniFrac distances (Table C.2).

**Effects of heat shock on microbial diversity.** In the "heat shock" rig, when comparing the distal tap microbiome pre- vs post- heat shock, alpha diversity indicated no obvious change either in terms of richness (Chao 1 index, Fig. C.2) in bulk water phase (immediately pre- vs immediately post-, P = 0.976; immediately pre- vs 8-hr post-, P = 0.991) or evenness (Gini index, Fig. C.3) in either biofilm (2-month pre- vs 2-month post-, P < 0.001, Gini index range 0.993-0.999 indicating limited difference) or bulk water (immediately pre- vs immediately post-, P = 0.928; immediately pre- vs 8-hr post-, P = 1.000) phases. The decrease of richness (Chao 1 index) in the biofilm phase after the heat shock treatment (Fig. C.2) was no greater than that observed in influent biofilm (1695, 1438, 1015 for 2-month pre-, immediately pre- and 2-month

post-, respectively in influent biofilm), which suggests that temporal factors were the more plausible driver than the heat shock treatment. The limited effect of heat shock on microbial diversity was further corroborated by the near perfect correlation of the relative abundances of individual OTUs pre- and post- heat shock (Fig. C.4).

**Comparison of heat shock versus thermal disruption.** A key question is whether adjusting the water heater temperature directly shapes the microbial community structure and essentially overrides the influence of previous temperature regimes. To evaluate this, the "thermal disruption" and "heat shock" rigs were compared after 2 months of maintaining both rigs at 40 °C following the heat shock. While the distal tap microbial communities were still distinct between the two rigs (both phases), the differences were smaller than those observed when comparing the pre- heat shock conditions, trending towards convergence (Table 4.2). Indeed, the distal tap bulk water microbiota from the two rigs remained distinct immediately ($R^2_{Adonis}$ = 0.7772, $P_{Adonis}$ = 0.001) and 8-hour post- heat shock ($R^2_{Adonis}$ = 0.6692, $P_{Adonis}$ = 0.001), where the *Bacteroidetes* phylum showed much lower relative abundance in the "thermal disruption" rig than in the "heat shock" rig (Fig 4.2).

**Table 4.2. Statistical comparison on the relative impact of rig, pipe orientation and water use frequency.** Weighted UniFrac distance matrix was applied for Adonis and Betadisp test with permutation = 999 ({vegan}, Oksanen et al., 2016, R).

| **a. Biofilm samples** | **Rig** | | | **Pipe Orientation** | | | **Water Use Frequency** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **$R^2_{Adonis}$** | **$P_{Adonis}$** | **$P_{Betadisp}$** | **$R^2_{Adonis}$** | **$P_{Adonis}$** | **$P_{Betadisp}$** | **$R^2_{Adonis}$** | **$P_{Adonis}$** | **$P_{Betadisp}$** |
| 2-month pre- (36) | 0.2917 | 0.001 | 0.261 | 0.1160 | 0.004 | 0.487 | 0.1508 | 0.006 | 0.290 |
| Immediately pre- (35) | 0.2082 | 0.001 | 0.014 | 0.1782 | 0.001 | 0.001 | 0.1897 | 0.003 | 0.873 |
| 2-month post- (36) | 0.1614 | 0.001 | 0.001 | 0.1925 | 0.001 | 0.024 | 0.1121 | 0.031 | 0.383 |
| **b.Bulk water samples** | **Rig** | | | **Pipe Orientation** | | | **Water Use Frequency** | | |
| | **$R^2_{Adonis}$** | **$P_{Adonis}$** | **$P_{Betadisp}$** | **$R^2_{Adonis}$** | **$P_{Adonis}$** | **$P_{Betadisp}$** | **$R^2_{Adonis}$** | **$P_{Adonis}$** | **$P_{Betadisp}$** |
| 2-month pre- (36) | 0.4311 | 0.001 | 0.012 | 0.1580 | 0.010 | 0.308 | 0.1018 | 0.104 | 0.466 |
| Immediately pre- (34) | 0.2937 | 0.001 | 0.001 | 0.0944 | 0.016 | 0.225 | 0.1125 | 0.056 | 0.282 |
| Immediately post- (36) | 0.7772 | 0.001 | 0.833 | 0.0116 | 0.645 | 0.714 | 0.0388 | 0.550 | 0.935 |
| 8-hour post- (36) | 0.6692 | 0.001 | 0.833 | 0.0253 | 0.364 | 0.965 | 0.0337 | 0.655 | 0.770 |
| 2-month post- (36) | 0.1076 | 0.013 | 0.896 | 0.0343 | 0.298 | 0.909 | 0.5350 | 0.001 | 0.051 |

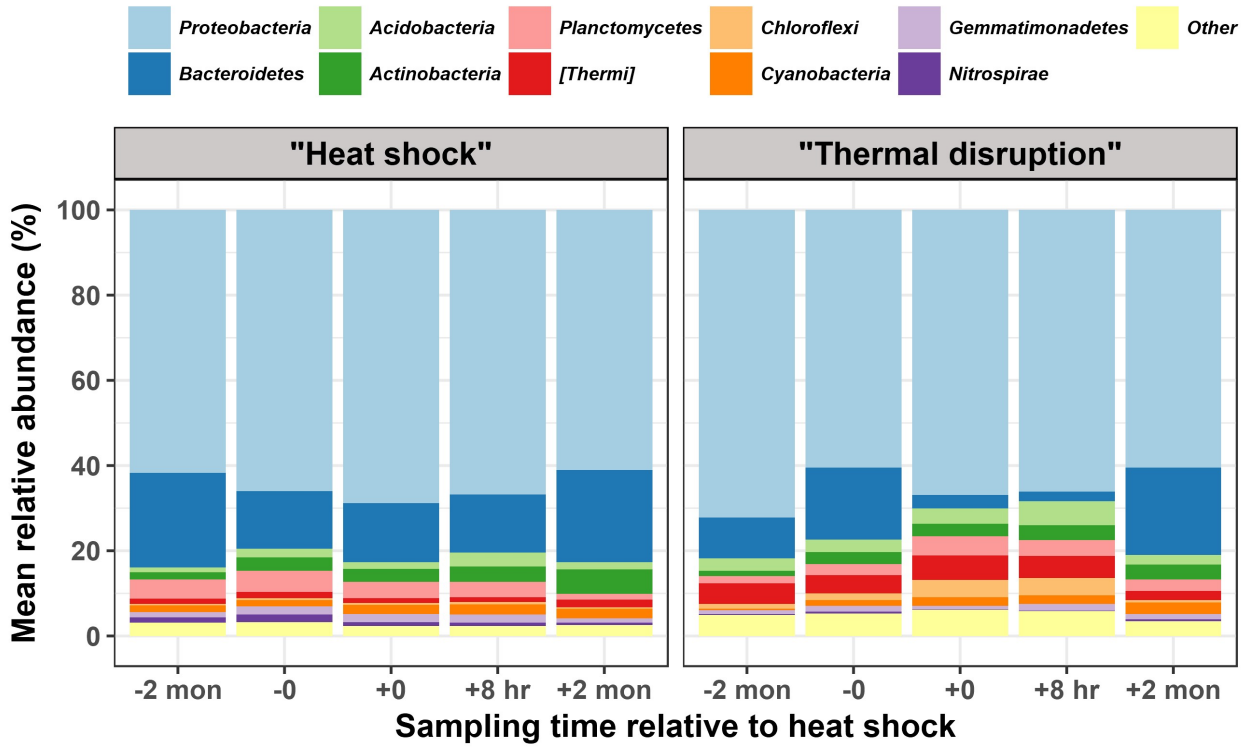


**Fig. 4.2. Top 10 abundant phyla in distal tap water samples across sampling time.** Abundant phyla were determined based on all samples combined and ranked in descending order by relative abundance from the top to the bottom of each bar in the figure. Remaining phyla were aggregated as "other" and always at the bottom of each bar.

**Influence of water heater temperature set point, pipe orientation, and water use frequency.** The relative importance of rig (representing water heater temperature set point difference), pipe orientation, and water use frequency on microbial community composition, as represented by dissimilarity matrices, varied with sampling time and phase (Table 4.2). For instance, at 2-month post- heat shock (both rigs at 40 °C), pipe orientation was the dominant factor for the biofilm phase, but had no significant effect for the bulk water phase (Table 4.2). Instead, water use frequency was the dominant factor shaping the bulk water microbial community composition. There also appeared to be potential synergistic factors between pipe orientation and water use frequency. For example, downward pipe orientation magnified the influence of water use frequency on the microbial composition of the biofilm (Fig. C.5; weighted UniFrac, downward n = 54, $R^2_{Adonis}$ = 0.2878, $P_{Adonis}$ = 0.001 vs upward n = 53, $R^2_{Adonis}$ = 0.0870, $P_{Adonis}$ = 0.019), likely by altering relative abundance more than occurrence of individual OTUs

(unweighted UniFrac, downward $R^2_{Adonis}$ = 0.0743, $P_{Adonis}$ = 0.003 vs upward $R^2_{Adonis}$ = 0.0661, $P_{Adonis}$ = 0.002). No such obvious interactions between pipe orientation and water use frequency were observed in bulk water phase.

Notably, among the "heat shock" rig distal tap biofilm samples (n = 53), pipe orientation ($R^2_{Adonis}$ = 0.2378, $P_{Adonis}$ = 0.001) and water use frequency ($R^2_{Adonis}$ = 0.2026, $P_{Adonis}$ = 0.001) appeared to be more influential than heat shock (pre- vs post- heat shock, $R^2_{Adonis}$ = 0.0662, $P_{Adonis}$ = 0.01), although all three factors indicated small effects when considering the unweighted UniFrac distance matrix ($R^2_{Adonis}$ = 0.0574/ 0.0826/ 0.0877, $P_{Adonis}$ = 0.001).

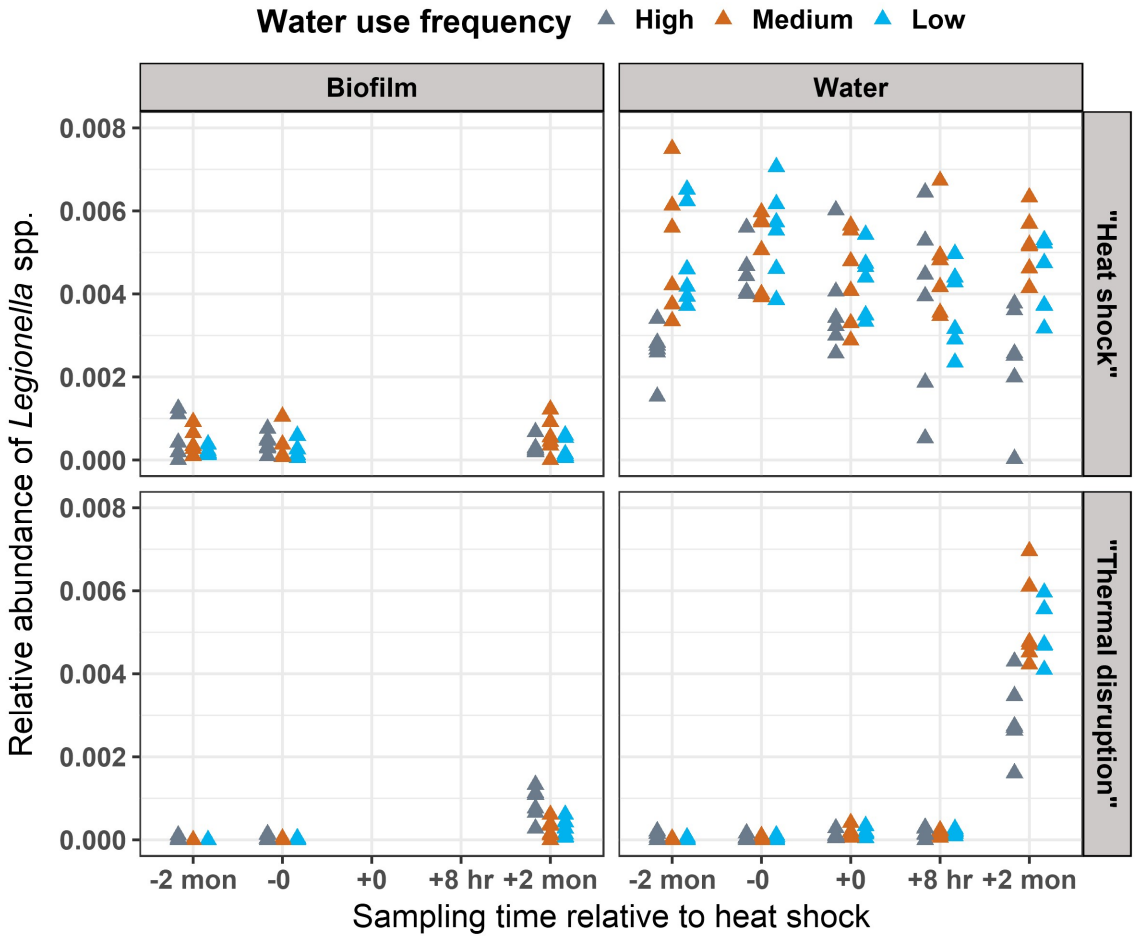


**Fig. 4.3. Relative abundance of *Legionella* spp. in distal tap samples.** Relative abundance was calculated from 16S rRNA gene amplicon sequencing data.

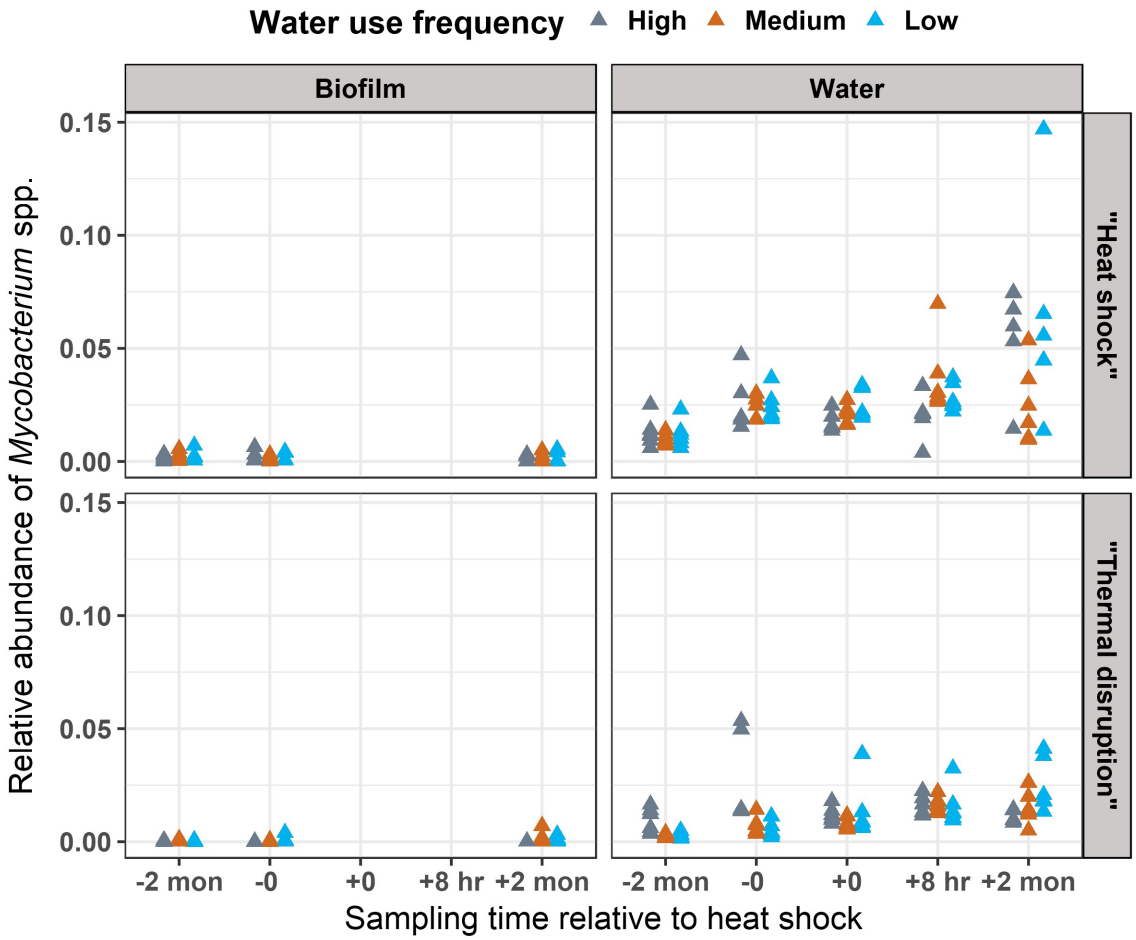


**Fig. 4.4. Relative abundance of *Mycobacterium* spp. in distal tap samples.** Relative abundance data was calculated from 16S rRNA gene amplicon sequencing data.

**Impact of heat shock on *Legionella* spp. and *Mycobacterium* spp. based on 16S rRNA amplicon sequencing.** The relative abundance of *Legionella* spp. at the distal tap remained relatively constant in both the biofilm and bulk water across all time points in the "heat shock" rig. By contrast, *Legionella* spp. relative abundance increased markedly in both phases of the "thermal disruption" rig following operation for 2 months at the reduced temperature of 40 °C (both phases, Fig. 4.3). High water use frequency (shortest stagnation period) was associated with low relative abundance of *Legionella* spp. in pre- and post- heat shock bulk water samples from the "heat shock" rig, and post- heat shock bulk water samples from the "thermal disruption" rig. Interestingly, *Mycobacterium* spp. were enriched post- heat shock in the bulk water phase of the "heat shock" rig (both for high and low water use frequency, Fig. 4.4; no obvious trend in

influent). *Mycobacterium* spp. did not appear to be as sensitive to water heater temperature set point as *Legionella* spp. (e.g., 60 °C pre-heat shock vs 40° post-heat shock samples in the "thermal disruption" rig). The "thermal disruption" rig was also associated with lower planktonic *Mycobacterium* relative abundance relative to the "heat shock" rig, suggesting that longer term operation at 60 °C may have kept levels lower, even after dropping to 40 °C for two months.

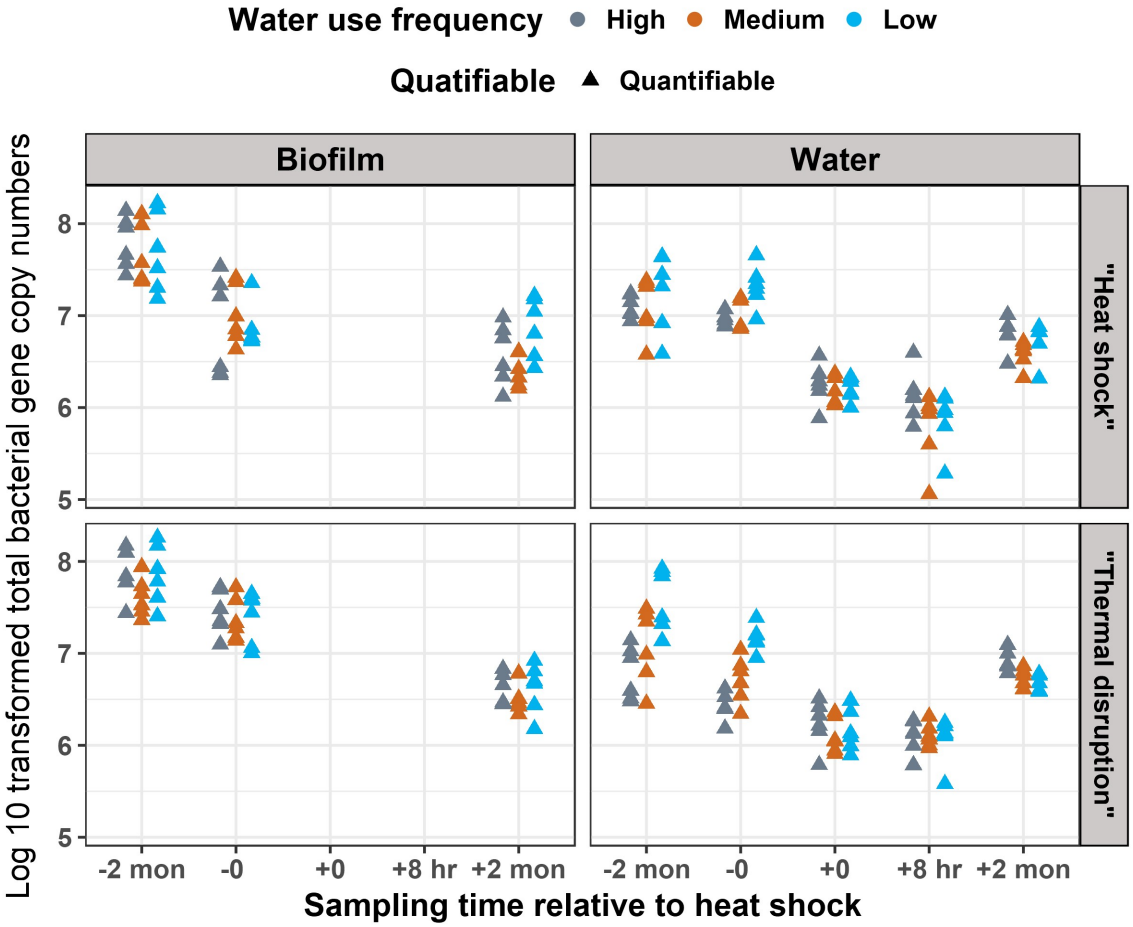


**Fig. 4.5. Total bacterial gene copy numbers in distal tap samples by qPCR.** The gene copy numbers were log 10 transformed.

**Impact of heat shock on total bacteria, *L. pneumophila*, *V. vermiformis*, and *M. avium* gene copy numbers by qPCR.** Quantitative PCR measurements indicated lowest total bacteria numbers at 2-month post- heat shock in both rigs in both biofilm and bulk water (Fig. 4.5). In the "heat shock" rig, increased relative abundance (normalized to the 16S rRNA gene copy

numbers) of *L. pneumophila* was observed in the biofilm following heat shock (Fig. 4.6). It should be noted, however, that *L. pneumophila* relative abundance was variable in the bulk water during the two baseline samplings prior to the heat shock (2-month pre- and immediately pre-heat shock). Relative abundance of *L. pneumophila* increased in the bulk water 2-month post-heat shock relative to immediately pre- heat shock, but was comparable to levels 2-month pre-heat shock.

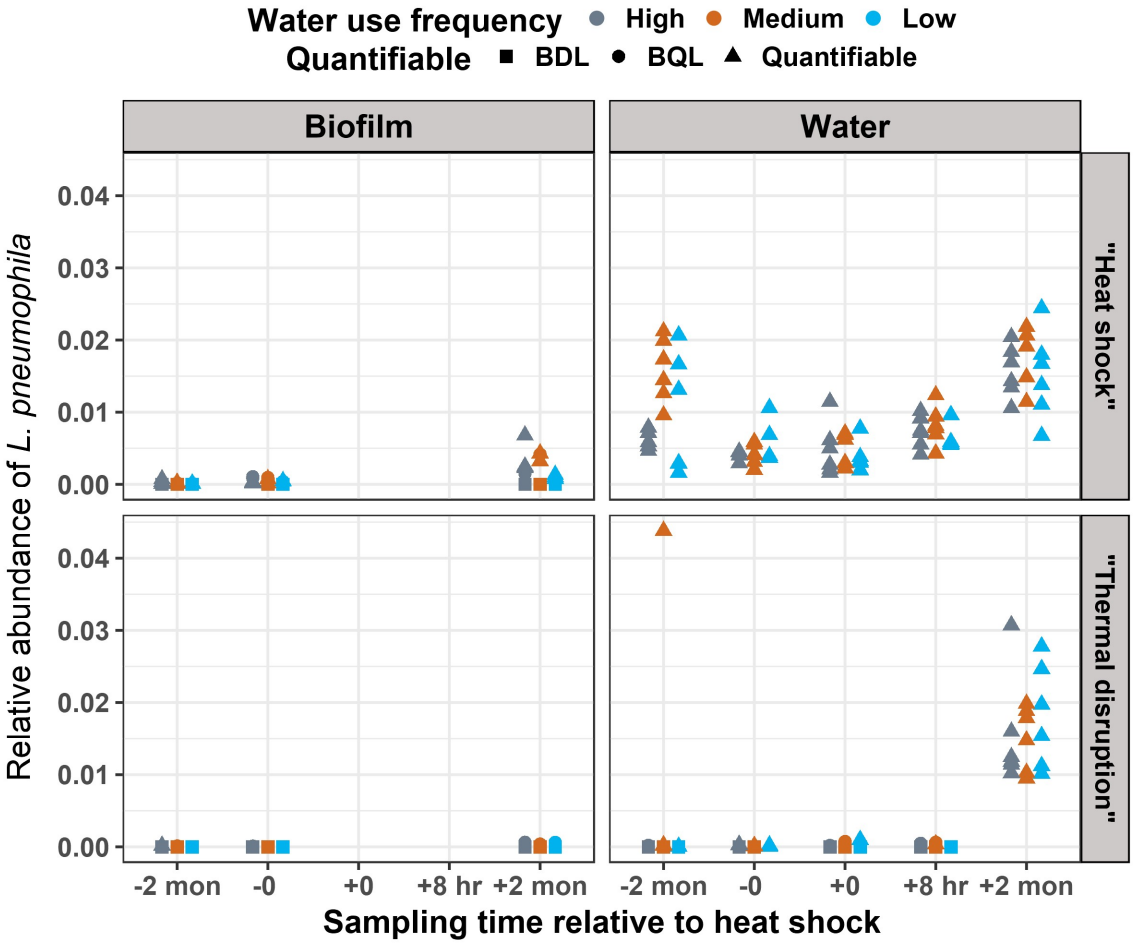


**Fig. 4.6. Relative abundance of *Legionella pneumophila* in distal tap samples by qPCR.**

In the "thermal disruption" rig, decreased water heater temperature had no apparent effect on relative abundance of *L. pneumophila* in the biofilm of distal taps, with consistently low levels observed pre- and post- heat shock (Fig. 4.6). However, in the bulk water, the relative abundance of *L. pneumophila* increased substantially 2-month after lowering the temperature to

40 °C, to a level comparable to those observed after 2 months at 40 °C post- heat shock in the "heat shock" rig (Fig. 4.6).

Post- heat shock samples tended to harbor higher relative abundance of *V. vermiformis* compared to pre- heat shock samples (both phases, both rigs, Fig. 4.7), although the opposite trend was observed in absolute numbers of *V. vermiformis* genes in the biofilm of the "heat shock" rig (Fig. C.6). In the "heat shock" rig, distal tap bulk water achieved peak relative abundance of *V. vermiformis* at 8-hour post- heat shock, which is likely due to lowest total bacterial gene copy numbers (Fig. 4.5) rather than corresponding peak *V. vermiformis* gene copy numbers (Fig. C.6). Interestingly, in post- heat shock samples, the water use frequency associated with lowest relative abundance of *V. vermiformis* seemed to bifurcate between biofilm (low water use frequency) and bulk water (high water use frequency) conditions.

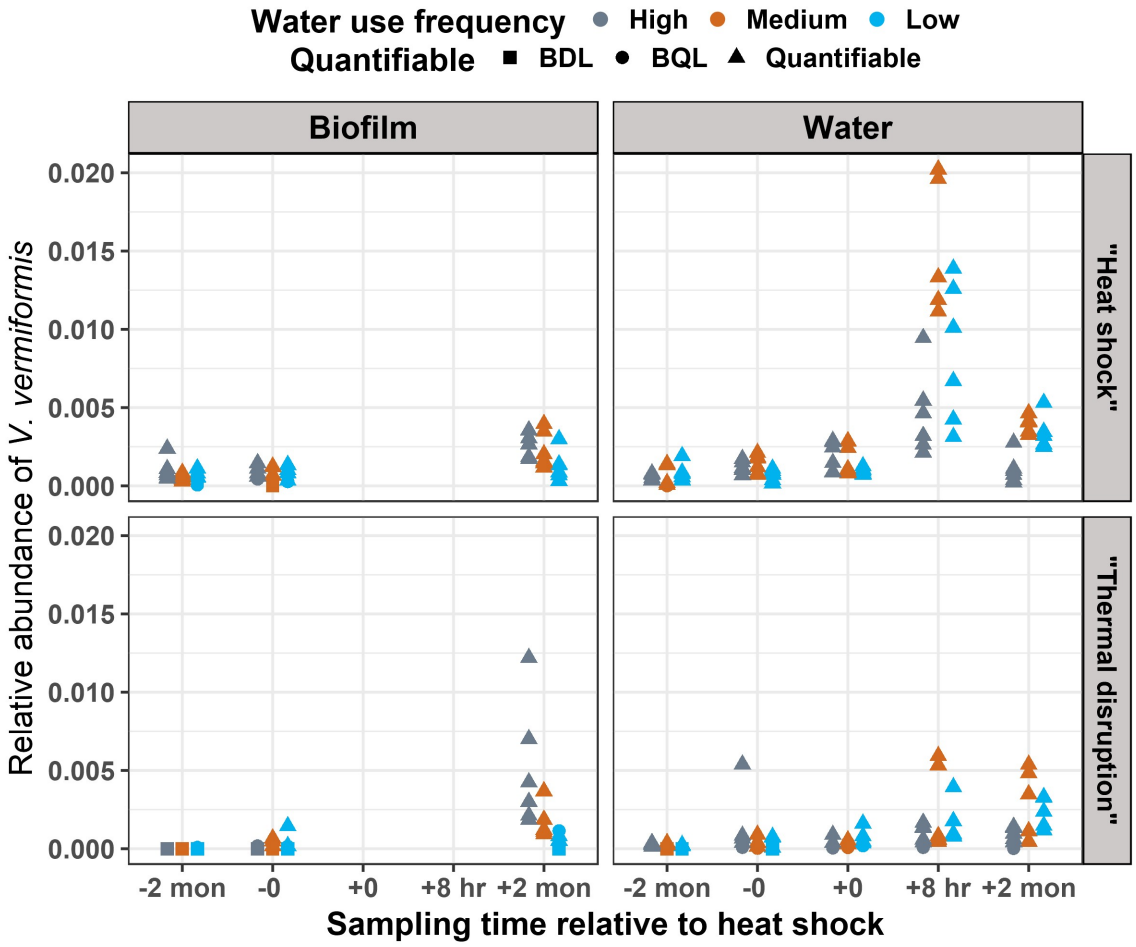


**Fig. 4.7. Relative abundance of *Vermamoeba vermiformis* in distal tap samples by qPCR.**

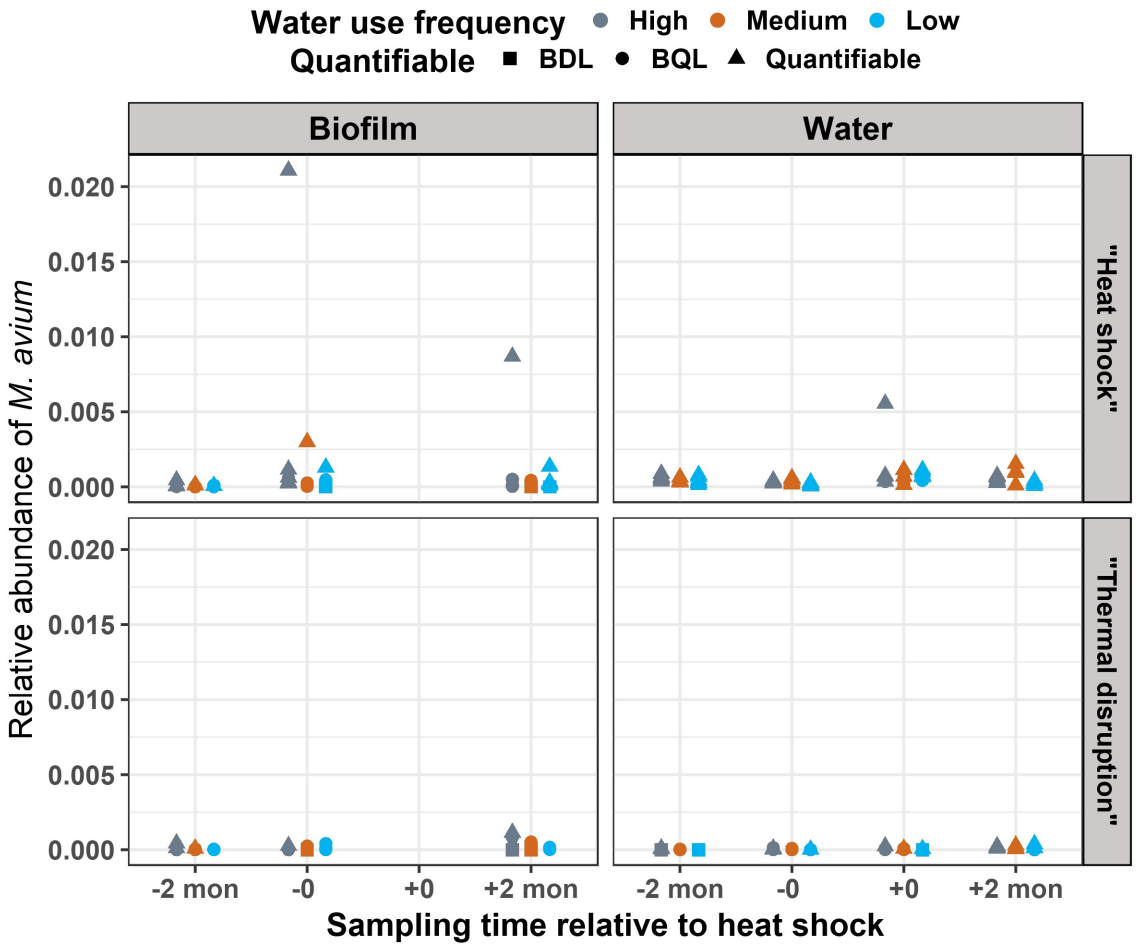


**Fig. 4.8. Relative abundance of *Mycobacterium avium* in distal tap samples by qPCR.** Samples at 8-hr post- heat shock were not included.

Relative abundance of *M. avium* was consistently low (< 0.5%) in both phases of both rigs, except for three high 'outlier' values and indicated no obvious trend with respect to heat shock (Fig. 4.8). The high 'outlier' values appeared to be a mixed result of decreased total bacteria numbers (ranking 241, 249, and 180 out of a total of 249 samples in descending order) and/or elevated *M. avium* numbers (ranking 20, 66, and 4 out of a total of 249 samples in descending order; Figure C.7). The top two highest 'outlier' values were from biofilm samples collected at the same distal tap (downward oriented, high water use frequency pipe in the "heat shock" rig, at immediately pre- and 2-month post- heat shock). The third highest 'outlier' value was from a water sample collected at a distal tap (upward oriented, high water use frequency pipe in the "heat shock" rig, at immediately post- heat shock) with high *M. avium* gene copy numbers noted

pre- heat shock. Notably, all three 'outlier' values were associated with the "heat shock" rig and high water use frequencies, likely suggesting potential conditions which would enrich *M. avium*. Combining data from the 40 °C conditions, 91 distal tap samples had quantifiable *L. pneumophila* and *V. vermiformis* numbers. Spearman correlation analysis indicated a positive association between *L. pneumophila* and *V. vermiformis* numbers in the biofilm (Spearman's rank correlation coefficient = 0.48), while a negative association was noted in the bulk water (Spearman's rank correlation coefficient = -0.19, Fig. 4.9), though it was not visually apparent.

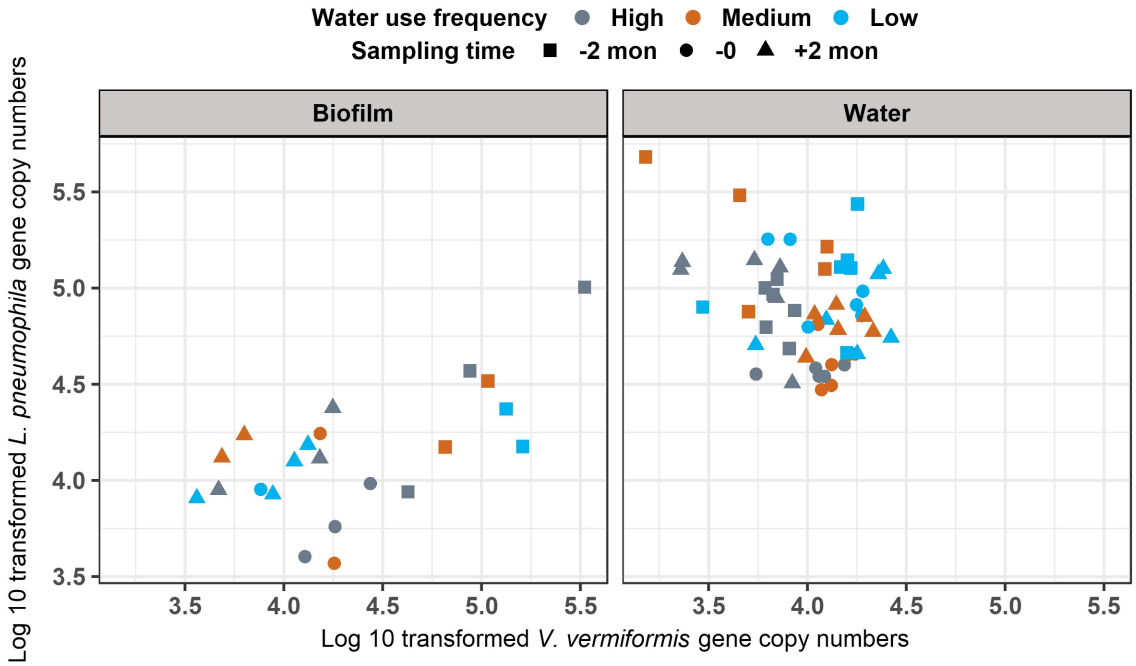


**Fig. 4.9. Relationship between *Legionella pneumophila* gene copy number and *Vermamoeba vermiformis* gene copy number.** Samples included for analysis were: (1) distal tap samples, with (2) water heater temperature set point at 40 °C, and (3) collected at 2-month pre-, immediately pre- and 2-month post- heat shock, who had (4) quantifiable *L. pneumophila* and *V. vermiformis* numbers by qPCR. Both gene copy numbers were log transformed with base 10.

## DISCUSSION

Heat shock alone, as employed in this study at the mild end of the spectrum of published methodologies, did not substantially influence the distal tap microbiota in terms of richness (Chao 1 index), evenness (Gini index), OTU counts, or microbial community dissimilarity

patterns. The consistent trends based both on weighted and unweighted UniFrac distance matrices further demonstrated no apparent bias of heat shock affect towards abundant versus rare OTUs. At 2-month post- heat shock, relative abundance of OPs and associated amoeba hosts either remained constant or increased slightly compared to pre- heat shock levels. Relative to the heat shock procedure as applied in the current study, water heater temperature set point and water use frequency each appeared to be more influential in terms of microbial community composition and *Legionella* control.

While from a microbial perspective a sudden increase by 20 °C is certainly a "shock," the null effect observed on biofilm microbiota could possibly be attributed to microbial resistance or resilience. The failure of the heat shock to disturb established biofilm microbiota is consistent with what has been called the "resistance" or insusceptibility scenario (Bridier et al., 2011), suggesting ineffectiveness of the applied heat shock procedure as a microbiota control measure. Alternatively, in what might be called a "resilience" scenario, the microbial composition in the biofilm phase was indeed shifted by the heat shock, yet recovered close to pre- heat shock status within the subsequent 2 months. In this scenario, heat shock could effectively control biofilm microbiota if applied at an appropriate frequency. A previous study using a pilot-scale hot water distribution system (Farhat et al., 2012) found that heat shock (water heater temperature set point at 70 °C and flushing of all taps for 30 min) conferred the most transitory effect on microbiota in the biofilm phase. In 1 of the 2 heat shock events where change in microbial composition was observed post heat shock, the change was only observed on days 1, 3 and 7, sharing highly similar biofilm microbial composition to that of the non-heat-treatment conditions. Interestingly, in the present study where repeated swabbing of the same surface area for biofilm collection was employed with time, the composition of the biofilm was still highly stable in spite of different temperature regimes. This further underscores the conclusion that biofilms were relatively stable in the current study.

Unlike the relatively 'stationary' biofilm phase where the influence of prior temperature regimes may be more long-lasting, the impact of the heat shock on the bulk water phase is expected to be more 'transient'. When comparing bulk water immediately pre- and immediately post- heat shock in the "heat shock" rig, direct impacts were noted in terms of shifts in relative abundance,

rather than occurrences of individual OTUs (Table C.2; Fig. 4.2). At 2-month post- heat shock, changes in bulk water microbial composition were minor and comparable to natural background fluctuation prior to the heat shock. With a simplified model of assuming complete mixing in the water heater and recirculating line, while ignoring microbial activity (i.e., growth and death), the bulk water affected by heat shock should be washed out of the system and thus not represent direct effects at 2-month post- heat shock (the proportion of water directly impacted by heat shock remained $< 1\times10^{-15}$, Table C.3). Indeed, microbes favored by heat shock theoretically would be subject to 'wash out' (i.e., water use) following loss of their thermophilic niche. Still, long-term influence of heat shock on the bulk water phase is possible via seeding of the bulk water from biofilm during stagnation, to the extent the biofilm phase was impacted in the longer-term. However, this scenario is not as likely in this study given that there was limited influence of heat shock on the biofilm microbiota.

In this study, water heater temperature set point (60 vs 40 °C) and water use frequency appeared to be more influential than heat shock alone, both in terms of effect on the microbial community composition and control of *Legionella*. Ideally, a heat shock procedure achieves a sufficiently high water heater temperature set point (e.g., as high as > 70 °C) to exert a lethal effect on microbes throughout the system under continuous flow conditions, whereas the present study focuses on the mild end of the heat shock spectrum and may represent a 'worst case scenario' of lack of observable benefit. Bédard et al. (2016a) applied a thermal disinfection procedure by maintaining the temperature setting above 70 °C for 1 hour and flushing at the point-of-use for > 7 minutes in a 400-bed university hospital in Canada. They found that in one of the two hot water systems examined, thermal disinfection failed to affect the planktonic culturable *L. pneumophila* levels, where enhanced long-term thermal regime was a more important control measure than thermal disinfection. From a microbial ecology point of view, we speculate that the ineffectiveness of heat shock on *L. pneumophila* control is associated with its inability to affect distal tap microbial communities, even 2-month post- heat shock. Despite the absence of a uniform trend, all three high 'outlier' values of *M. avium* relative abundance (by qPCR) were associated with high water use frequency and the "heat shock" rig, which was largely maintained at an inferior water heater temperature set point of 40 °C. It has been hypothesized in a previous study that higher water use frequency delivers more nutrients to distal taps within the same

period, which can encourage OP growth under conditions of low water heater temperature set points and no disinfectant residual (Rhoads et al., 2017).

It should be emphasized that heat shock might unintentionally exacerbate OP exposure. First, we suspect that heat shock could induce a competitive advantage for OPs in the long-term, even if they are reduced in the short-term, as in the case of heat-pretreatment to select for more *Legionella* versus other bacteria prior to culturing (CDC, 2005). Consistent with this hypothesis, heat shock did effectively decrease the total bacteria numbers, which could have contributed to lower OPs levels. However, *L. pneumophila* tended to recover quicker (Fig. C.8) than total bacteria (Fig. 4.5), and become more enriched with time, especially in the bulk water phase ("heat shock" rig, Fig. 4.6). This is consistent with the findings of a stagnant-water model study (Vervaeren et al., 2006), in which heat-treated tap water (60 °C for 30 min, subsequently cooled down to room temperature) contributed to elevated *L. pneumophila* levels in both biofilm and water phases compared to untreated tap water. It is of future research interest to determine how long it takes the total bacteria to be restored to pre- heat shock levels and whether the relative abundance of *L. pneumophila* will remain elevated during this process (i.e., increased *L. pneumophila* exposure risk). *V. vermoformis* was enriched post- heat shock mainly due to decreased total bacteria numbers (Fig. 4.5) rather than decreased *V. vermiformis* gene copy numbers (Fig. C.6), potentially as an indicator of their feeding on the bacteria. *V. vermiformis* and other amoebae are known for their crucial role in *L. pneumophila*'s life cycle (Kuiper et al., 2004), higher relative abundance of which would have likely contributed to the increased levels of *L. pneumophila*. Second, heat shock might not be a sustainable method for engineering control of the microbiota: as each application would likely enhance the resistance of biofilm microbiota to the standard heat shock procedure, similar to the case of drought on the tropical forest soil microbial communities (Bouskill et al., 2013). In addition, repeated heat shock operations might select for thermophilic *L. pneumophila* (able to grow above 50 °C, Lesnik et al., 2016) or lead to thermo-acclimated *Legionella* post- heat shock (Farhat et al., 2010).

Given that the heat shock treatment applied in this study represented the mild end of published guidelines, it would be of future interest to investigate a full range of heat shock procedures. While setting the water heater temperature at > 70 °C and flushing distal taps for 30 minutes

would be expected to be more effective for OP control, this is not known for certain and it is plausible that selection of OPs could actually be stronger. In reality, the heat shock procedure a homeowner could reasonably achieve is to set the water heater temperature at the highest level: typically, 65.5 °C for residential electric water heaters (AO Smith. Technical Bulletins, Bulletin 31) and 71.1 °C for residential gas water heaters (AO Smith. Technical Bulletins, Bulletin 32). Also, in large buildings with complex plumbing systems, it can be especially difficult to achieve and maintain target at-the-tap flushing temperatures. It would also be of future research interest to explore how the functional profiles of microbial communities are affected by heat shock protocols, as the present study largely focused on taxonomic responses.

## CONCLUSIONS

The study evaluates the effect of heat shock as a microbiota and *Legionella* control measure. The mild heat shock procedure adopted in this study, i.e., maintaining at-the-tap flushing temperatures above 55 °C for approximately 30 min with water heater temperature set point at 60 °C, conferred little change in biofilm and bulk water microbiota at distal taps, where water is used and exposure occurs. Importantly and consistent with prior research, increased relative abundances of *L. pneumophila* and *V. vermiformis* were observed post- heat shock. Water heater temperature set point and water use frequency appeared to be promising long-term alternatives in terms of microbiota modification and *L. pneumophila* control.

## DECLARATIONS

**Ethics approval and consent to participate:** Not applicable.

**Consent for publication:** Not applicable.

**Availability of data and materials:** The datasets generated and/or analyzed during the current study are available in the QIITA database under Study ID 10504 (https://qiita.ucsd.edu/) and European Nucleotide Archive (ENA) under accession number PRJEB22241.

**Competing interests:** The authors declare that they have no competing interests.

**Funding:** The project was financed by The Alfred P. Sloan MoBE program, the National Science Foundation (CBET Award #1336650) and the Institute for Critical Technology and Applied Science (ICTAS) at Virginia Tech.

**Authors' contributions:** PJ performed the 16S rRNA gene sequencing, data analysis, generated the figures, and wrote the manuscript. WJR was in charge of building and maintaining the system, collecting the samples, performing DNA extraction and qPCR experiments, and tabulated raw qPCR results. WJR, MAE, AP, and PJ conceived and designed the experiments. WJR, MAE, AP revised and edited the draft. All authors read and approved the final manuscript.

**Acknowledgements:** The authors would like to thank all undergraduates (Kelly Peterson, Arjan Ahluwalia, Daniel Logsdon) for maintaining the system, assisting sample collection and DNA extraction; Christina Swann for conducting qPCR experiments.

## REFERENCES

Anderson MJ. (2001). A new method for non-parametric multivariate analysis of variance. *Aust Ecol* 26: 32-46. doi: 10.1111/j.1442-9993.2001.01070.pp.x.

Anderson MJ. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* 62: 245-253. doi: 10.1111/j.1541-0420.2005.00440.x.

AO Smith. Technical Bulletins, Bulletin 31: Temperature adjustment (residential electric). http://www.hotwater.com/lit/bulletin/bulletin31.pdf. Accessed 10 August 2017.

AO Smith. Technical Bulletins, Bulletin 32: Temperature adjustment (residential gas). http://www.hotwater.com/lit/bulletin/bulletin32.pdf. Accessed 10 August 2017.

Association of Water Technologies (AWT). (2003). *Legionella* 2003: An Update and Statement by the Association of Water Technologies. Association of Water Technologies. http://www.legionella.com/images/awtlegionella2003.pdf. Accessed 22 May 2017.

Baron JL, Harris JK, Holinger EP, Duda S, Stevens MJ, Robertson CE, et al. (2015). Effect of monochloramine treatment on the microbial ecology of *Legionella pneumophila* and

associated bacterial populations in a hospital hot water system. *Syst Appl Microbiol* 38: 198 – 205. doi: 10.1016/j.syapm.2015.02.006.

Bédard E, Boppe I, Koumamé S, Martin P, Pinsonneault L, Valiquette L, et al. (2016a). Combination of heat shock and enhanced thermal regime to control the growth of a persistent *Legionella pneumophila* strain. *Pathogens* 5(2): 35. doi:10.3390/pathogens5020035.

Bédard E, Lévesque S, Martin P, Pinsonneault L, Paranjape K, Lalancette C, et al. (2016b). Energy conservation and the promotion of *Legionella pneumophila* growth: the probable role of heat exchangers in a nosocomial outbreak. *Infec Control Hosp Epidemiol* 37: 1475-1480. doi: 10.1017/ice.2016.205.

Bouskill NJ, Lim HC, Borglin S, Slave R, Wood TE, Silver WL, et al. (2013). Pre-exposure to drought increases the resistance of tropical forest soil bacterial communities to extended drought. *ISME J* 7: 384 – 394. doi:10.1038/ismej.2012.113.

Bridier A, Briandet R, Thomas V, Dubois-Brissonnet F. (2011). Resistance of bacterial biofilms to disinfectants: a review. *Biofouling* 27: 1017-1032. doi: 10.1080/08927014.2011.626899.

Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R. (2010a). PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26(2): 266-267. doi: 10.1093/bioinformatics/btp636.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al. (2010b). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7: 335. doi: 10.1038/nmeth.f.303.

CDC. (2005). Procedures for the recovery of *Legionella* from the environment. https://www.cdc.gov/legionella/health-depts/inv-tools-cluster/lab-inv-tools/procedures-manual.pdf. Accessed 23 May 2017.

Darelid J, Löfgren S, Malmvall BE. (2002). Control of nosocomial Legionnaires' disease by keeping the circulating hot water temperature above 55 °C: experience from a 10-year surveillance programme in a district general hospital. *J Hosp Infect* 50(3): 213-219. doi: 10.1053/jhin.2002.1185.

Declerck P. (2010). Biofilms: the environmental playground of *Legionella pneumophila*. *Environ Microbiol* 2: 557 – 566. doi: 10.1111/j/1462-2920.2009.02025.x.

Douterelo I, Sharpe RL, Boxall JB. (2013). Influence of hydraulic regimes on bacterial community structure and composition in an experimental drinking water distribution system. *Water Res* 47(2): 503 – 516. doi: 10.1016/j.watres.2012.09.053.

Earth Microbiome Project. http://www.earthmicrobiome.org/. Accessed 2 April 2017.

Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19): 2460 – 2461. doi: 10.1093/bioinformatics/btq461.

EPA. (2016). National Primary Drinking Water Regulations (NPDWR). https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulation-table. Accessed 21 May 2017.

Falkinham III JO, Pruden A, Edwards M. (2015). Opportunistic premise plumbing pathogens: increasingly important pathogens in drinking water. *Pathogens* 4: 373 – 386. doi: 10.3390/pathogens4020373.

Farhat M, Trouilhé M-C, Briand E, Moletta-Denat M, Robine E, Frère J. (2010). Development of a pilot-scale 1 for *Legionella* elimination in biofilm in hot water network: heat shock treatment evaluation. *J Appl Microbiol* 108(3): 1073 – 1082. doi: 10.1111/j.1365-2672.2099.04541.x.

Farhat M, Moletta-Denat M, Frère J, Onillon S, Trouilhé M-C, Robine E. (2012). Effects of disinfection of *Legionella* spp., eukarya, and biofilms in a hot water system. *Appl Environ Microbiol* 78(19): 6850 – 6858. doi: 10.1128/AEM.00831-12.

Hua F, West JR, Barker RA, Forster CF. (1999). Modeling of chlorine decay in municipal water supplies. *Water Res* 33: 2735 – 2746. doi: 10.1016/S0043-1354.

Ji P, Parks JL, Edwards MA, Pruden A. (2015). Impact of water chemistry, pipe material and stagnation on the building plumbing microbiome. *PLoS ONE* 10(10): e0141087. 10.1371/journal.pone.0141087.

Ji P, Rhoads WJ, Edwards MA, Pruden A. (2017). Impact of water heater temperature setting and water use frequency on the building plumbing microbiome. *ISME J* 11: 1318 – 1330. doi: 10.1038/ismej.2017.14.

Kuiper MW, Wullings BA, Akkermans ADL, Beumer RR, van der Kooij D. (2004). Intracellular proliferation of *Legionella pneumophila* in *Hartmannella vermiformis* in aquatic biofilms grown on plasticized polyvinyl chloride. *Appl Environ Microbiol* 70: 6826 – 6833. 10.1128/AEM.70.11.6826-6833.2004.

Kuiper MW, Valster RM, Wullings BA, Boonstra H, Smidt H, van der Kooij D. (2006). Quantitative detection of the free-living amoeba *Hartmannella vermiformis* in surface water by using real-time PCR. *Appl Environ Microbiol* 72: 5750 – 5756. doi: 10.1128/AEM.00085-06.

Lautenschlager K, Boon N, Wang Y, Egli T, Hammes F. (2010). Overnight stagnation of drinking water in household taps induces microbial growth and changes in community composition. *Water Res* 44(17): 4868 – 4877. doi: 10.1016/j.watres.2010.07.032.

Lesnik R, Brettar I, Höfle MG. (2016). *Legionella* species diversity and dynamics from surface reservoir to tap water: from cold adaptation to thermophily. *ISME J* 10: 1064 – 1080. doi:10.1038/ismej.2015.199.

Lozupone C, Knight R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71: 8228 – 8235. doi: 10.1128/AEM.71.12.8228-8235.2005.

Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. (2012). PANDAseq: Paired-eND Assembler for Illumina sequences. *BMC Bioinformatics* 13: 31. doi: 10.1186/1471-2105-13-31.

McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A et al. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 6: 610 – 618. doi: 10.1038/ismej.2011.139.

Nazarian EJ, Bopp DJ, Saylors A, Limberger RJ, Musser KA. (2008). Design and implementation of a protocol for the detection of Legionella in clinical and environmental samples. Diagn *Microbiol Infect Dis* 62: 125 – 132. doi: 10.1016/j.diagmicrobio.2008.05.004.

Nguyen C, Elfland C, Edwards M. (2012). Impact of advanced water conservation features and new copper pipe on rapid chloramine decay and microbial regrowth. *Water Res* 46: 611 – 620. doi: 10.1016/j.watres.2011.11.006.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB et al. (2016). Vegan: Community Ecology Package. R package version 2.3-5. https://cran.r-project.org/web/packages/vegan/vegan.pdf. Accessed 2 April 2017.

Pinto AJ, Schroeder J, Lunn M, Sloan W, Raskin L. (2014). Spatial-temporal survey and occupancy-abundance modeling to predict bacterial community dynamics in the drinking water microbiome. *Mbio* 5(3): e01135 – 14. doi: 10.1128/mBio.01135-14.

Pohlert T. (2016). The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR). R package version 4.1. 2016. http://cran.r-project.org/web/packages/PMCMR/PMCMR.pdf.

Pozos N, Scow K, Wuertz S, Darby J. (2004). UV disinfection in a model distribution system: biofilm growth and microbial community. *Water Res* 38: 3083 – 3091. doi: 10.1016/j.watres.2004.04.011.

Price MN, Dehal PS, Arkin AP. (2010). FastTree2 – Approximately Maximum-Likelihood trees for large alignments. *PLoS ONE* 5(3): e9490. doi: 10.1371/journal.pone.0009490.

R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. Accessed 22 May 2017.

Rhoads WJ, Pruden A, Edwards MA. (2014). Anticipating challenges with in-building disinfection for control of opportunistic pathogens. *Water Environ Res* 86(6): 540 – 549. doi: 10.2175/106143014X13975035524989.

Rhoads WJ, Ji P, Pruden A, and Edwards MA. (2015). Water heater temperature set point and water use patterns influence *Legionella pneumophila* and associated microorganisms at the tap. *Microbiome* 3: 67. doi: 10.1186/s40168-015-0134-1.

Rhoads WJ, Pruden A, Edwards MA. (2016). Convective mixing in distal pipes exacerbates *Legionella pneumophila* growth in hot water plumbing. *Pathogens* 5(1): 29. doi: 10.3390/pathogens5010029.

Rhoads WJ, Pruden A, Edwards MA. (2017). Interactive effects of corrosion, copper, and chloramines on *Legionella* and mycobacteria in hot water plumbing. *Environ Sci Technol* 51(12): 7065 – 7075. doi: 10.1021/acs.est.6b05616.

Stout JE, Best MG, Yu VL. (1986). Susceptibility of members of the family Legionellaceae to thermal stress: implications for heat eradication methods in water distribution systems. *Appl Environ Microbiol* 52: 396-399.

Stout JE, Yu VL. (2003). Experiences of the first 16 hospitals using copper-silver ionization for *Legionella* control: implications for the evaluation of other disinfection modalities. *Infec Control Hosp Epidemiol* 24(8): 563 – 568.

Suzuki MT, Taylor LT, DeLong EF. (2000). Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5 '-nuclease assays. *Appl Environ Microbiol* 66: 4605 – 4614. doi: 10.1128/AEM.66.11.4605-4614.2000.

Vervaeren H, Temmerman R, Devos L, Boon N, Verstraete W. (2006). Introduction of a boost of *Legionella pneumophila* into a stagnant-water model by heat treatment. *FEMS Microbiol Ecol* 58(3): 583 – 592. doi: 10.1111/j.1574-6941.2006.00181.x.

Wang H, Edwards MA, Falkinham III JO, Pruden A. (2013). Probiotic Approach to Pathogen Control in Premise Plumbing Systems? A Review. *Environ Sci Technol* 47: 10117 – 10128. doi: 10.1021/es402455r.

Wang Q, Garrity GM, Tiedje JM, Cole JR. (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73(16): 5261 – 5267. doi: 10.1128/AEM.0062-07.

Ware GW. (1989). Control of *Legionella* in plumbing systems. *Rev Environ Contam Toxicol* 107: 79 – 92.

Wang H, Edwards M, Falkinham III JO, 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. *Appl Environ Microbiol* 78: 6285 – 6294. doi: 10.1128/AEM.01492-12.

Wilton S, Cousins D. (1992). Detection and identification of multiple mycobacterial pathogens by DNA amplification in a single tube. *Genome Res* 1: 269 – 273. doi: 10.1101/gr.1.4.269.

Yu VL, Best M, Stout J, Brown A, Goetz A. Effectiveness of intermittent short term elevations of the hospital water supply in controlling nosocomial Legionnaires' disease [abstr L 19]. Presented at the 82nd meeting of the American Society for Microbiology, Atlanta, 7-12 March 1982.

## APPENDIX C – SUPPLEMENTARY INFORMATION FOR CHAPTER 4

### Table C.1. Summary of heat shock treatment procedures.

| Entity | Water heater | | Distal tap | | | Reference |
|---|---|---|---|---|---|---|
| | Temperature set point | Duration | Flush at water temperature of | Duration | Continuous or periodic flushing* | |
| **Florida DoH** | 71-77 °C | 24 hr | Not specified | 5 - 20 min | Continuous | [1] |
| **OSHA** | 70 °C | 24 hr | Not specified | 20 min | Continuous | [2] |
| **WHO** | Not specified | Not specified | 50 - 60 °C | Not specified | Periodic | [3] |
| **ASHRAE** | 71 - 77 °C | Not specified | Not specified | Not specified | Continuous | [4] |
| **AWT** | > 60 °C, preferably 66 °C | Not specified | Not specified | Up to 30 min | Continuous | [5] |
| **This study**** | 60 °C | | > 55 °C | 30 min | Periodic | |

* For water heater, some procedures recommend to set the water heater temperature at a higher number and maintain for a certain duration before flushing; all flushing operations are considered "continuous" unless the procedure specified "periodic" flushing. "Not specified" indicates no information.

** The heat treatment procedure adopted in this study was targeting at the AWT protocol with reference from Stout et al. 1986 [6]. Distal taps were flushed intermittently to achieve >55 °C water temperature for 30 min. The complete-mixed water heater tank + recirculating pipe holds 73.9 L water, while flushing rate was 3.8 L/min at each distal tap. If continuous flushing simultaneously at all distal taps was selected (total flushing rate of 68.4 L/min), at-the-tap flushing temperature cannot be maintained sufficiently high to inhibit Legionella. As an alternative, the periodic (intermittent) flushing was opted and distal pipes were flushed in sets.

[1] Florida Department of Health (Florida DOH). Guidelines for the Surveillance, Investigation, and Control of Legionnaires' Disease in Florida. 2014. http://www.floridahealth.gov/diseases-and-conditions/legionnaires-disease/_documents/gsi-legionella-update-final.pdf. Accessed 22 May 2017.
[2] Occupational Safety and Health Administration (OSHA). OSHA Technical Manual TED 01-00-015. 2014. https://www.osha.gov/dts/osta/otm/otm_iii/otm_iii_7.html#5. Accessed 22 May 2017.
[3] Bartram J, Chartier Y, Lee JV, Pond K, Surman-Lee S. Eds. Legionella and the Prevention of Legionellosis. 2007. http://www.who.int/water_sanitation_health/emerging/legionella.pdf. Accessed 22 May 2017.

[4] BSR/ASHRAE Standard 188P. Prevention of Legionellosis Associated with Building Water Systems. Second public review draft. 2011.
https://osr.ashrae.org/Public%20Review%20Draft%20Standards%20Lib/Std-188P-PPR2%20Final%206%2010%202011.pdf. Accessed 10 August 2017.
[5] Association of Water Technologies (AWT). Legionella 2003: An Update and Statement by the Association of Water Technologies. Association of Water Technologies. 2003.
https://www.awt.org/pub/035C2942-03BE-3BFF-08C3-4C686FB7395C. Accessed 10 August 2017.
[6] Stout JE, Best MG, Yu VL. Susceptibility of members of the family Legionellaceae to thermal stress: implications for heat eradication methods in water distribution systems. Appl Environ Microbiol. 1986; 52: 396-399.

**Table C.2. Distal tap water microbial dissimilarity patterns across sampling time.** Data recorded in the form of "$R^2_{Adonis}$ ($P_{Adonis}$); $P_{Betadisp}$" (permutation = 999).

| **a. "Heat shock" rig, weighted UniFrac** | (n = 96). Sampling time: 0.3404 (0.001); 0.001 | | | | |
|---|---|---|---|---|---|
| | **2-month pre- (20)** | **Immediately pre- (18)** | **Immediately post- (18)** | **8-hour post- (19)** | **2-month post- (21)** |
| **2-month pre- (20)** | | 0.2274 (0.001); 0.400 | 0.4381 (0.001); 0.001 | 0.3631 (0.001); 0.536 | 0.1434 (0.003); 0.018 |
| **Immediately pre- (18)** | | | 0.2006 (0.001); 0.001 | 0.1434 (0.001); 0.923 | 0.2209 (0.001); 0.006 |
| **Immediately post- (18)** | | | | 0.0774 (0.001); 0.002 | 0.2831 (0.001); 0.001 |
| **8-hour post- (19)** | | | | | 0.2420 (0.001); 0.006 |
| **2-month post- (21)** | | | | | |
| **b. "Heat shock" rig, unweighted UniFrac** | (n = 96). Sampling time: 0.1255 (0.001); 0.001 | | | | |
| | **2-month pre- (20)** | **Immediately pre- (18)** | **Immediately post- (18)** | **8-hour post- (19)** | **2-month post- (21)** |
| **2-month pre- (20)** | | 0.0598 (0.001); 0.604 | 0.1068 (0.001); 0.001 | 0.0955 (0.001); 0.312 | 0.0789 (0.001); 0.019 |
| **Immediately pre- (18)** | | | 0.0939 (0.001); 0.001 | 0.0806 (0.001); 0.561 | 0.0763 (0.001); 0.005 |
| **Immediately post- (18)** | | | | 0.0371 (0.001); 0.014 | 0.0989 (0.001); 0.001 |
| **8-hour post- (19)** | | | | | 0.0855 (0.001); 0.001 |
| **2-month post- (21)** | | | | | |
| **c. "Thermal disruption" rig, weighted UniFrac** | (n = 96). Sampling time: 0.47943 (0.001); 0.001 | | | | |
| | **2-month pre- (20)** | **Immediately pre- (18)** | **Immediately post- (18)** | **8-hour post- (19)** | **2-month post- (21)** |
| **2-month pre- (20)** | | 0.1038 (0.009); 0.471 | 0.23545 (0.001); 0.001 | 0.25766 (0.001); 0.001 | 0.38081 (0.001); 0.34 |
| **Immediately pre- (18)** | | | 0.31469 (0.001); 0.001 | 0.34156 (0.001); 0.001 | 0.24975 (0.001); 0.056 |
| **Immediately post- (18)** | | | | 0.08123 (0.002); 0.994 | 0.63108 (0.001); 0.002 |
| **8-hour post- (19)** | | | | | 0.64518 (0.001); 0.001 |
| **2-month post- (21)** | | | | | |
| **d. "Thermal disruption" rig, unweighted UniFrac** | (n = 96). Sampling time: 0.20185 (0.001); 0.001 | | | | |
| | **2-month pre- (20)** | **Immediately pre- (18)** | **Immediately post- (18)** | **8-hour post- (19)** | **2-month post- (21)** |
| **2-month pre- (20)** | | 0.04545 (0.003); 0.478 | 0.13205 (0.001); 0.001 | 0.12353 (0.001); 0.001 | 0.17374 (0.001); 0.499 |
| **Immediately pre- (18)** | | | 0.11542 (0.001); 0.001 | 0.1081 (0.001); 0.001 | 0.15927 (0.001); 0.979 |
| **Immediately post- (18)** | | | | 0.03159 (0.005); 0.257 | 0.20739 (0.001); 0.001 |
| **8-hour post- (19)** | | | | | 0.1956 (0.001); 0.001 |
| **2-month post- (21)** | | | | | |

**Table C.3. Simplified model for estimating proportion of directly impacted water along time.** "Remained" is the remaining "original" heat shocked water within water heater and recirculating line. Assumptions include: water heater and recirculating line is a complete mixed system with no (little) change between distal pipe and recirculating line; no microbial death or growth; no interference with biofilm; only considering physical mixing. Recirculating line sampling not accounted for.

| | | |
|---|---|---|
| **Tank + recirc line** | 71.90 | L |
| **Flush volume per distal tap** | 1.77 | L |

*3.8 L/min for 28 seconds*

| Sampling round | Time | % "Original" retained | "Original" retained/ L | Flushed volumn/ L | # of distal taps flushed | Distal taps flushed |
|---|---|---|---|---|---|---|
| **+0** | | 55.6% | 39.98 | 31.92 | 18 | All |
| **+ 8 hr** | | 30.9% | 22.23 | 31.92 | 18 | All |
| | 8 hr | 26.3% | 18.94 | 10.64 | 6 | High |
| | 16 hr | 22.4% | 16.14 | 10.64 | 6 | High |
| | 24 hr | 19.1% | 13.75 | 10.64 | 6 | High |
| | 32 hr | 16.3% | 11.72 | 10.64 | 6 | High |
| | 40 hr | 13.9% | 9.98 | 10.64 | 6 | High |
| | 48 hr | 11.8% | 8.50 | 10.64 | 6 | High |
| | 56 hr | 8.3% | 5.99 | 21.28 | 12 | High and medium |
| | 64 hr | 7.1% | 5.10 | 10.64 | 6 | High |
| | 72 hr | 6.0% | 4.35 | 10.64 | 6 | High |
| | 80 hr | 5.2% | 3.70 | 10.64 | 6 | High |
| | 88 hr | 4.4% | 3.16 | 10.64 | 6 | High |
| | 96 hr | 3.7% | 2.69 | 10.64 | 6 | High |
| | 104 hr | 3.2% | 2.29 | 10.64 | 6 | High |
| | 112 hr | 2.2% | 1.61 | 21.28 | 12 | High and medium |
| | 120 hr | 1.9% | 1.37 | 10.64 | 6 | High |
| | 128 hr | 1.6% | 1.17 | 10.64 | 6 | High |
| | 136 hr | 1.4% | 1.00 | 10.64 | 6 | High |
| | 144 hr | 1.2% | 0.85 | 10.64 | 6 | High |

| | | | | | |
|---|---|---|---|---|---|
| 152 hr | 1.0% | 0.72 | 10.64 | 6 | High |
| 160 hr | 0.9% | 0.62 | 10.64 | 6 | High |
| 168 hr | 0.5% | 0.34 | 31.92 | 18 | All |
| 176 hr | 0.4% | 0.29 | 10.64 | 6 | High |
| 184 hr | 0.3% | 0.25 | 10.64 | 6 | High |
| 192 hr | 0.3% | 0.21 | 10.64 | 6 | High |
| 200 hr | 0.3% | 0.18 | 10.64 | 6 | High |
| 208 hr | 0.2% | 0.15 | 10.64 | 6 | High |
| 216 hr | 0.2% | 0.13 | 10.64 | 6 | High |
| 224 hr | 0.1% | 0.09 | 21.28 | 12 | High and medium |
| 232 hr | 0.1% | 0.08 | 10.64 | 6 | High |
| 240 hr | 0.1% | 0.07 | 10.64 | 6 | High |
| 248 hr | 0.1% | 0.06 | 10.64 | 6 | High |
| 256 hr | 0.1% | 0.05 | 10.64 | 6 | High |
| 264 hr | 0.1% | 0.04 | 10.64 | 6 | High |
| 272 hr | 0.0% | 0.04 | 10.64 | 6 | High |
| 280 hr | 0.0% | 0.02 | 21.28 | 12 | High and medium |
| 288 hr | 0.0% | 0.02 | 10.64 | 6 | High |
| 296 hr | 0.0% | 0.02 | 10.64 | 6 | High |
| 304 hr | 0.0% | 0.02 | 10.64 | 6 | High |
| 312 hr | 0.0% | 0.01 | 10.64 | 6 | High |
| 320 hr | 0.0% | 0.01 | 10.64 | 6 | High |
| 328 hr | 0.0% | 0.01 | 10.64 | 6 | High |
| 336 hr | 0.0% | 0.01 | 31.92 | 18 | All |
| 344 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 352 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 360 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 368 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 376 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 384 hr | 0.0% | 0.00 | 10.64 | 6 | High |

| | | | | | |
|---|---|---|---|---|---|
| 392 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| 400 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 408 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 416 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 424 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 432 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 440 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 448 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| 456 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 464 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 472 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 480 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 488 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 496 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 504 hr | 0.0% | 0.00 | 31.92 | 18 | All |
| 512 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 520 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 528 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 536 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 544 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 552 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 560 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| 568 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 576 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 584 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 592 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 600 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 608 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 616 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| 624 hr | 0.0% | 0.00 | 10.64 | 6 | High |

| | | | | | |
|---|---|---|---|---|---|
| 632 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 640 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 648 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 656 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 664 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 672 hr | 0.0% | 0.00 | 31.92 | 18 | All |
| 680 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 688 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 696 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 704 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 712 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 720 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 728 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| 736 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 744 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 752 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 760 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 768 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 776 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 784 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| 792 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 800 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 808 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 816 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 824 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 832 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 840 hr | 0.0% | 0.00 | 31.92 | 18 | All |
| 848 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 856 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 864 hr | 0.0% | 0.00 | 10.64 | 6 | High |

| | | | | | |
|---|---|---|---|---|---|
| 872 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 880 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 888 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 896 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| 904 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 912 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 920 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 928 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 936 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 944 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 952 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| 960 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 968 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 976 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 984 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 992 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1000 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1008 hr | 0.0% | 0.00 | 31.92 | 18 | All |
| 1016 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1024 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1032 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1040 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1048 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1056 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1064 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| 1072 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1080 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1088 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1096 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| 1104 hr | 0.0% | 0.00 | 10.64 | 6 | High |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 1112 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1120 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| | 1128 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1136 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1144 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1152 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1160 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1168 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1176 hr | 0.0% | 0.00 | 31.92 | 18 | All |
| | 1184 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1192 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1200 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1208 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1216 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1224 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1232 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| | 1240 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1248 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1256 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1264 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1272 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1280 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1288 hr | 0.0% | 0.00 | 21.28 | 12 | High and medium |
| | 1296 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1304 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1312 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1320 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1328 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| | 1336 hr | 0.0% | 0.00 | 10.64 | 6 | High |
| **+ 2 mon** | 1344 hr | 9.93E-16 | 0.00 | 31.92 | 18 | All |

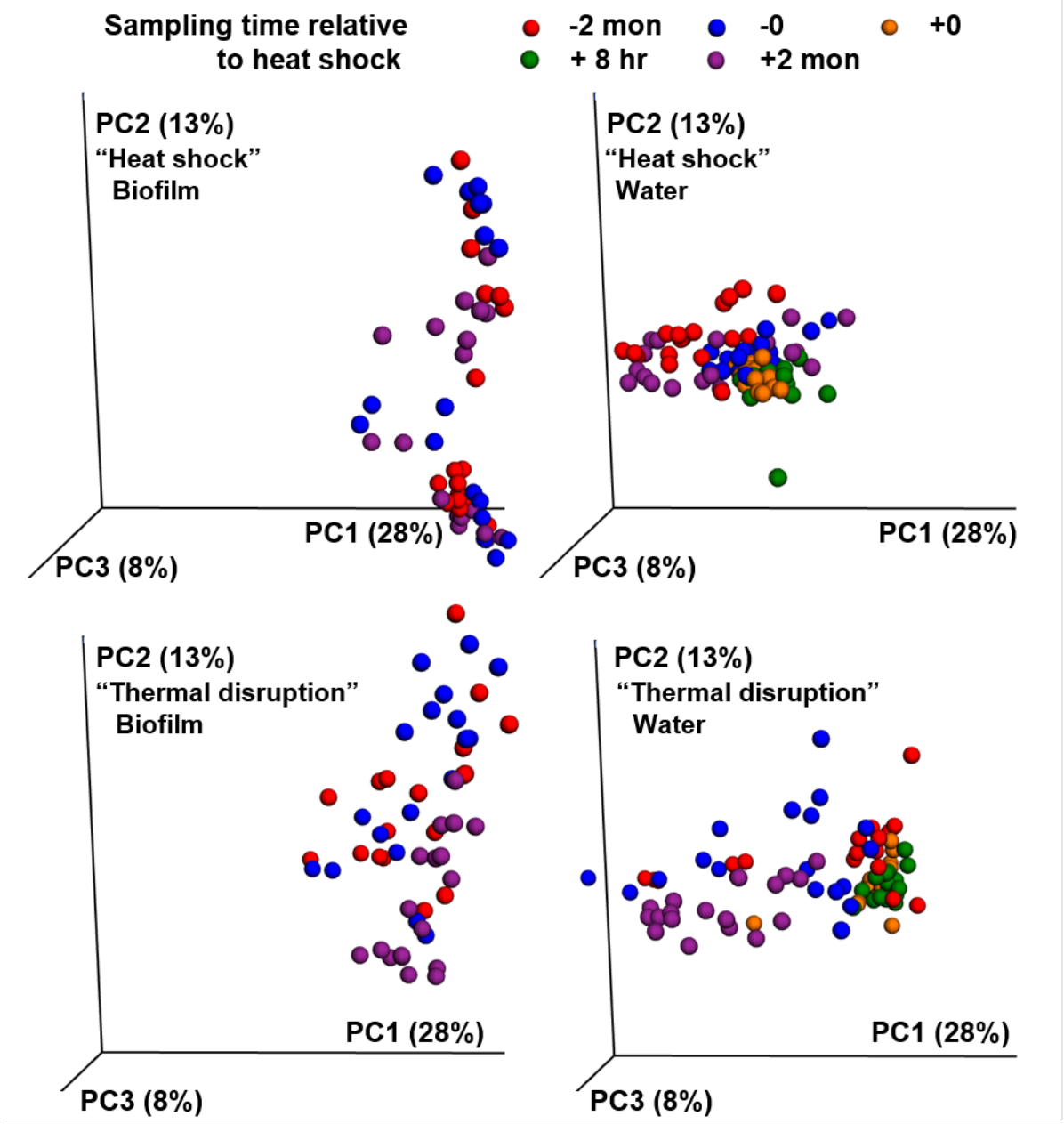


**Fig. C.1. Principal Coordinate Analysis on distal tap microbiome composition.** Figures are 3D Principal Coordinates Analysis based on weighted UniFrac distance matrices (rarefied to sequencing depth of 5, 200 for 100 times). Samples shown were distal tap ones.

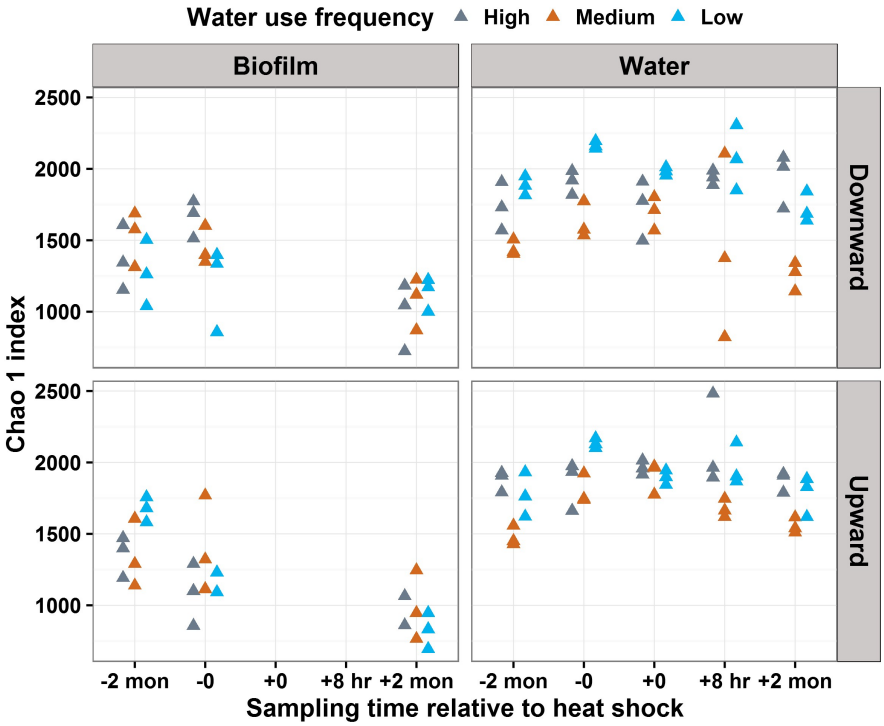


**Fig. C.2. Chao 1 index of "heat shock" rig distal tap samples across time.** Chao 1 index value is the average of 100 calculations based on rarefied OTU tables.

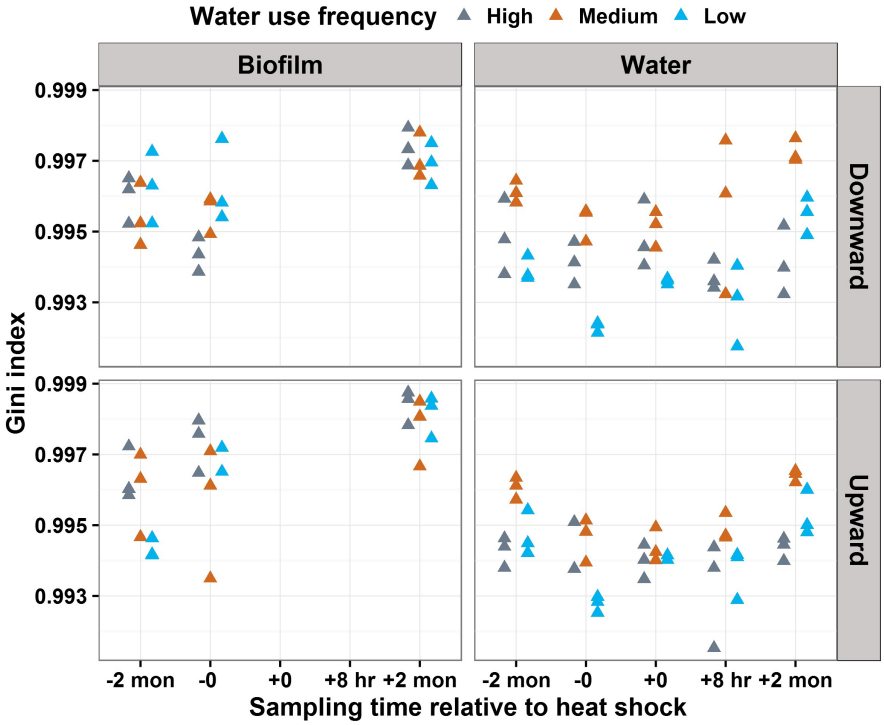


**Fig. C.3. Gini index of "heat shock" rig distal tap samples across time.** Gini index value is the average of 100 calculations based on rarefied OTU tables.

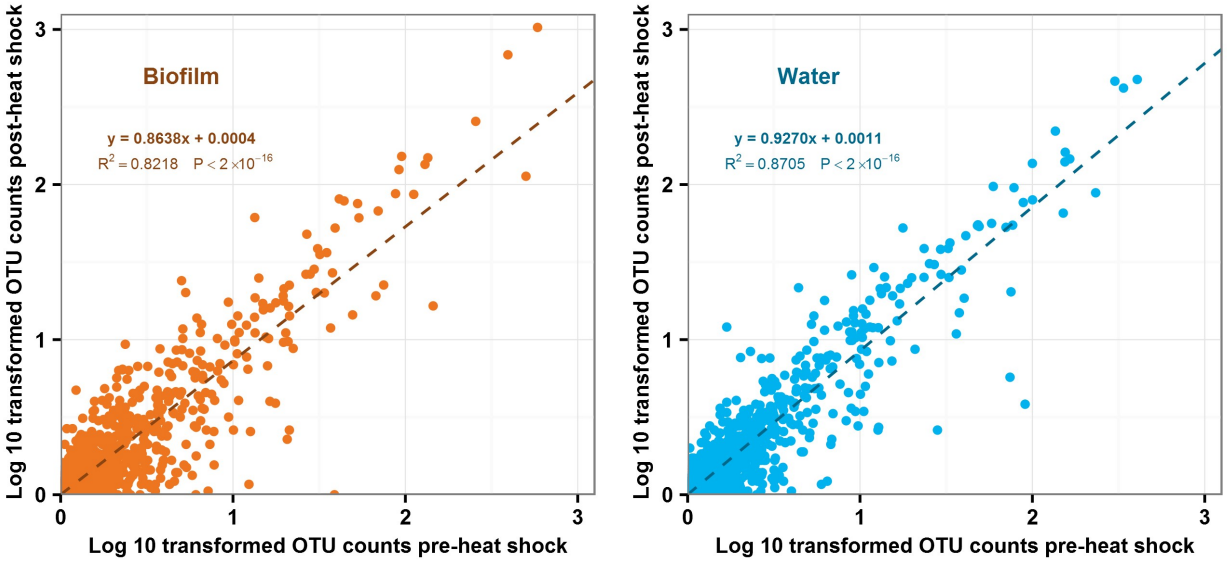


**Fig. C.4. OTU counts pre- and post- heat shock comparison.** All samples were from "heat shock" rig distal taps. Pre- heat shock included samples from 2-month pre- and immediately pre-heat shock, while post- heat shock included only samples from 2-month post- heat shock. Only OTUs detected at least once in either pre- or post- heat shock samples were included in this analysis. OTU counts (sequences per OTU) were first transformed as log10(OTU counts + 1). Linear regression was carried using transformed OTU counts.

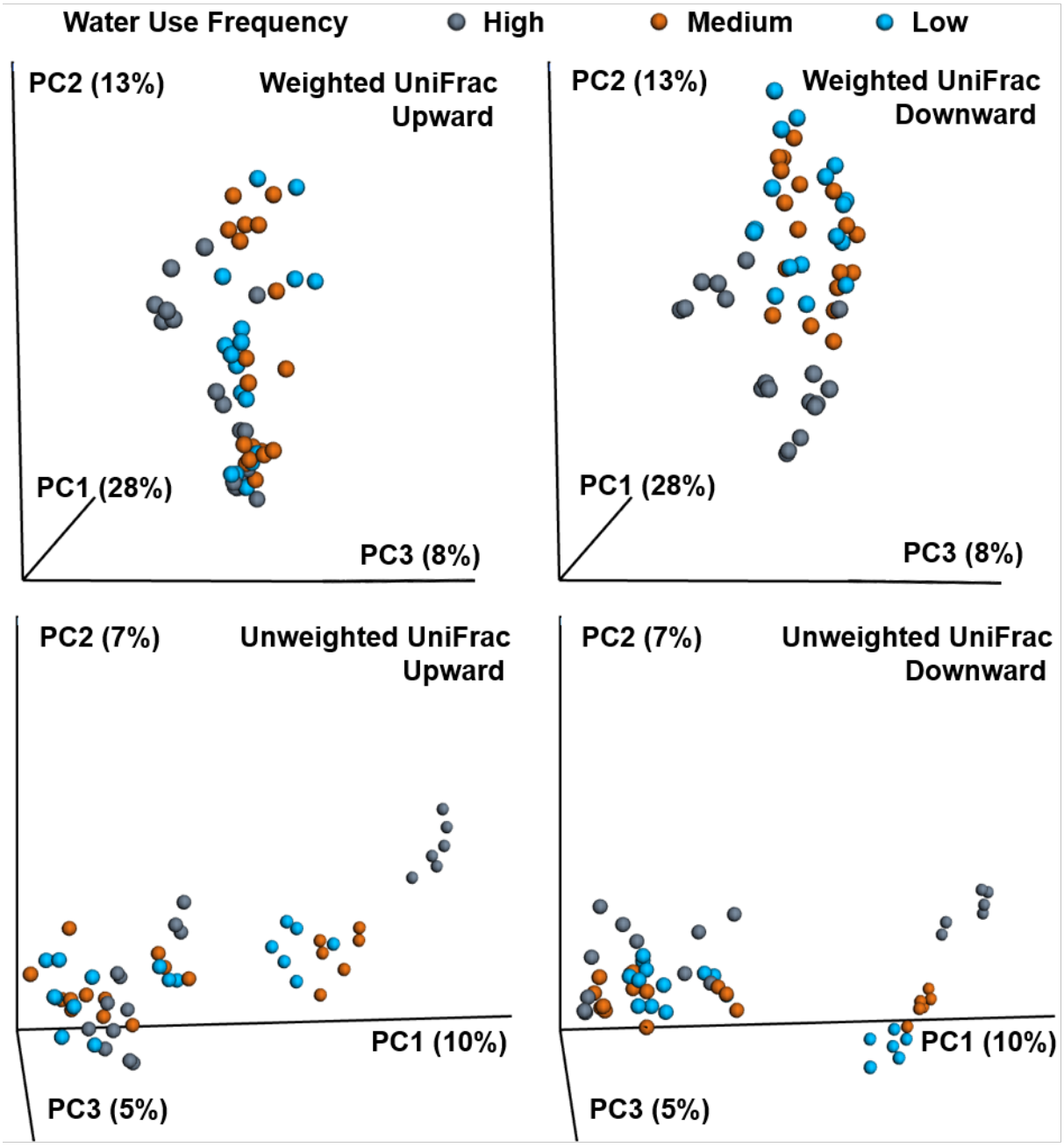


**Fig. C.5. Synergistic effect between pipe orientation and water use frequency.** Figures are 3D Principal Coordinates Analysis based on weighted (top row) and unweighted (bottom row) UniFrac distance matrices (rarefied to sequencing depth of 5, 200 for 100 times). Samples shown were distal tap biofilm ones.

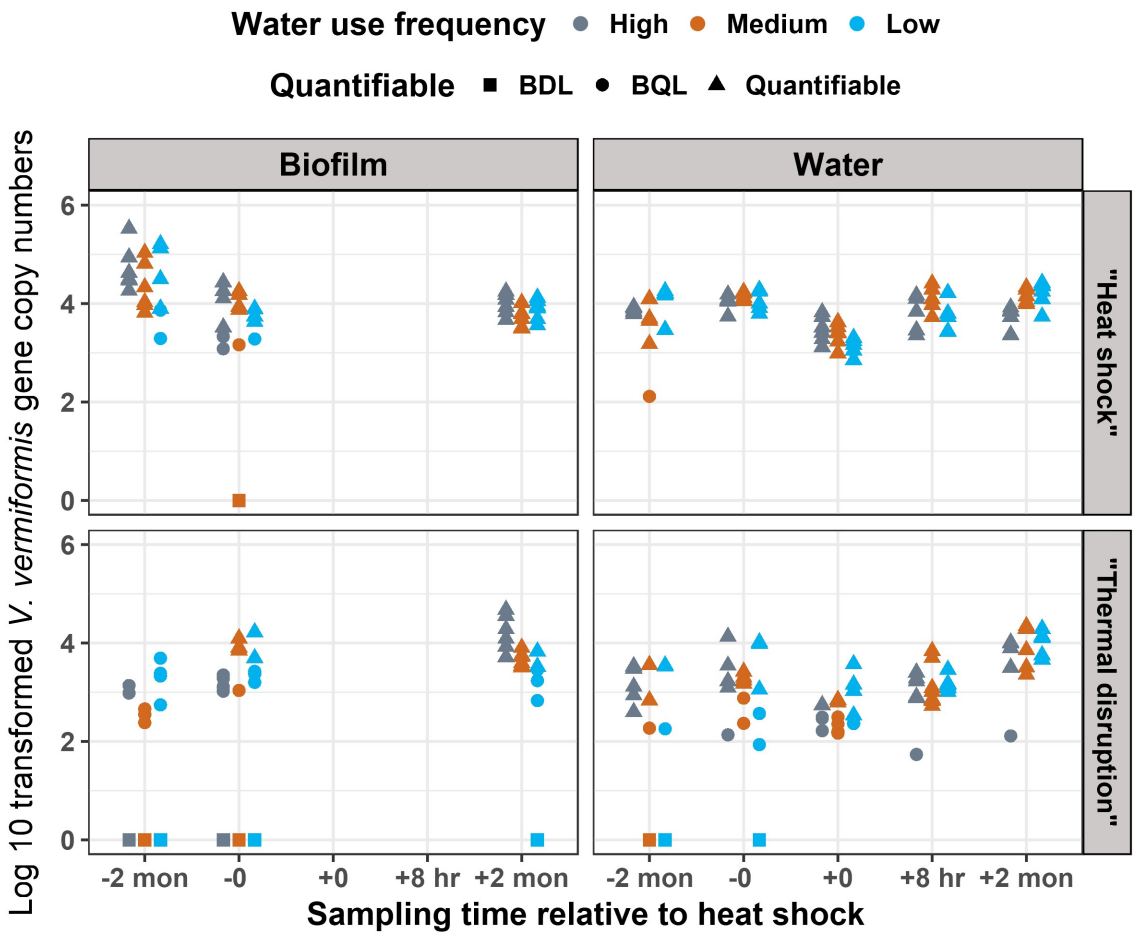


**Fig. C.6.** ***Vermamoeba vermiformis*** **gene copy numbers in distal tap samples by qPCR.** The gene copy numbers were log 10 transformed, i.e., gene copy number of X corresponds to $\log_{10}(X+1)$.

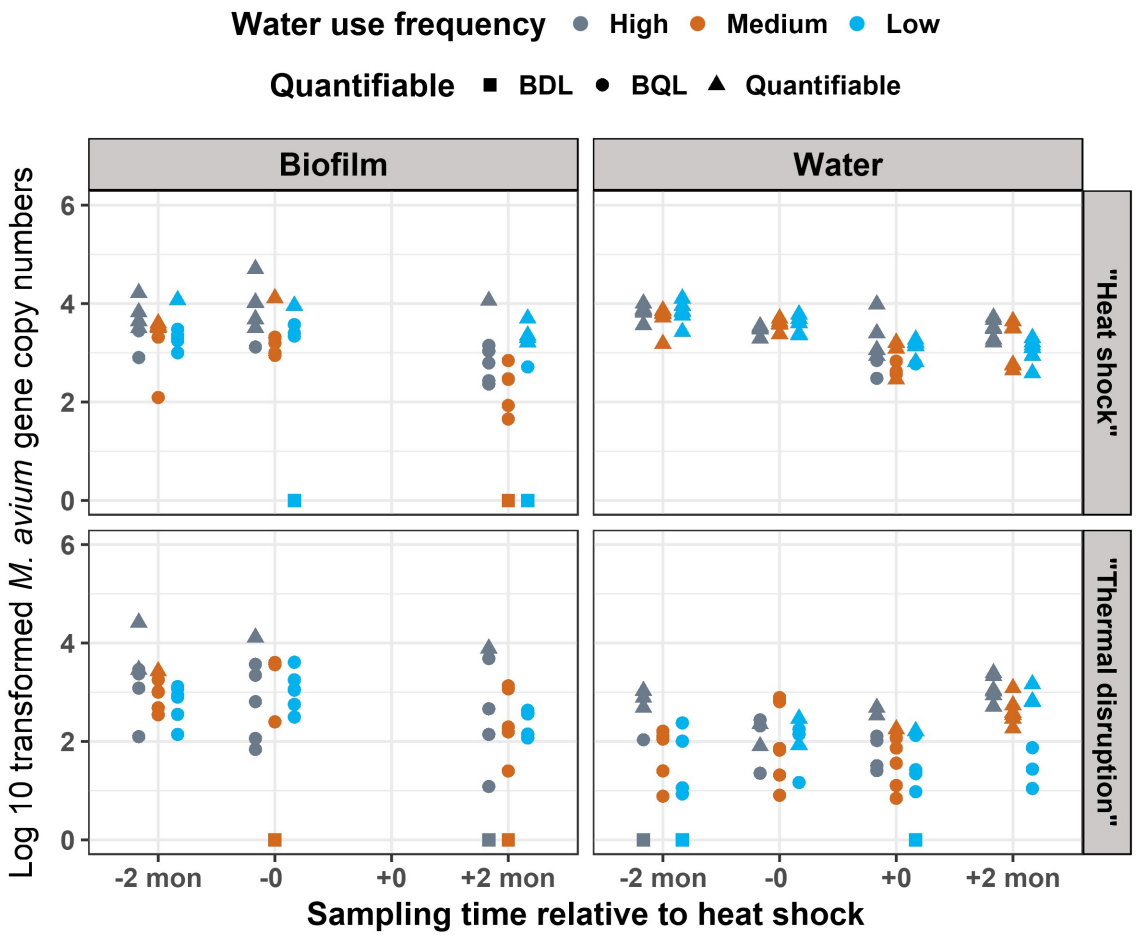


**Fig. C.7.** ***Mycobacterium avium*** **gene copy numbers in distal tap samples by qPCR.** Samples collected at 8-hour post- heat shock (+8 hr) was not included. The gene copy numbers were log 10 transformed, i.e., gene copy number of X corresponds to $\log_{10}(X+1)$.

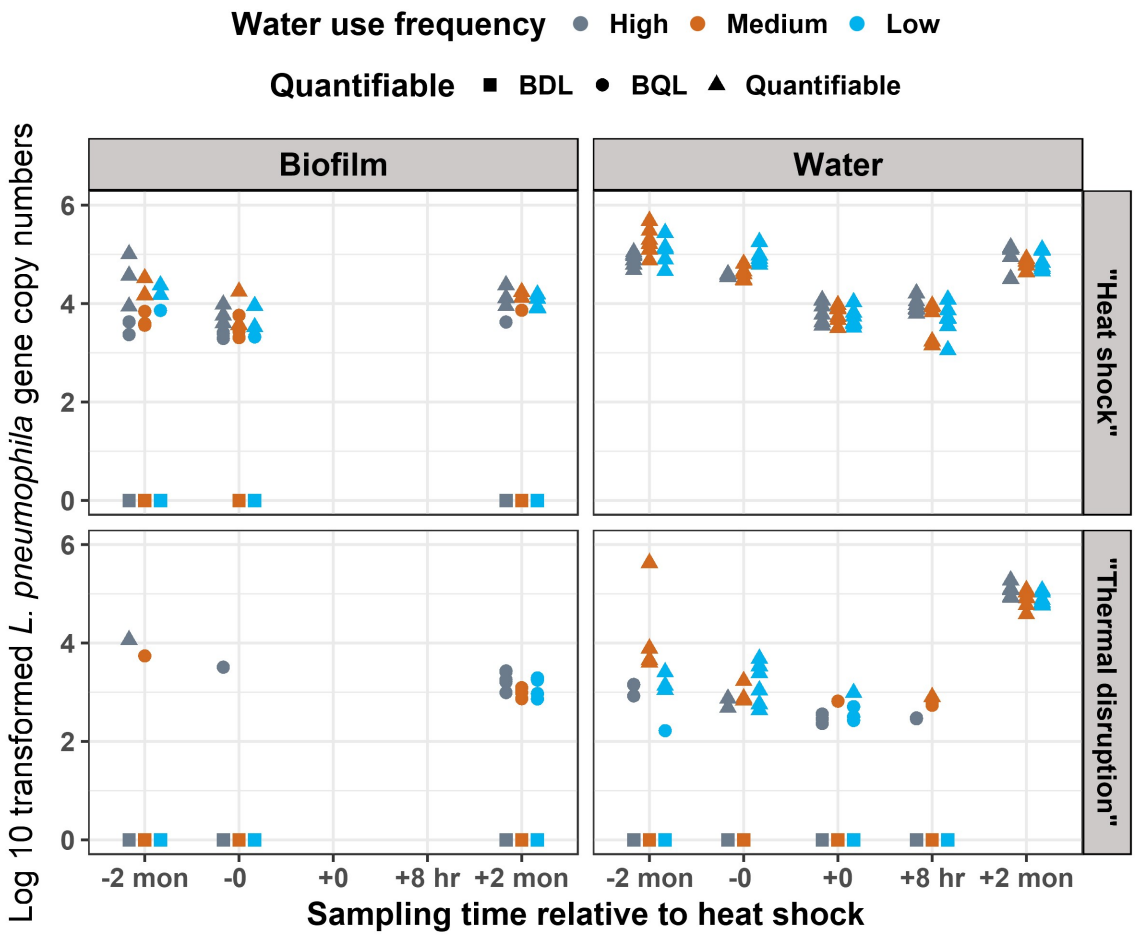


**Fig. C.8.** ***Legionella pneumophila*** **gene copy numbers in distal tap samples by qPCR.** The gene copy numbers were log 10 transformed, i.e., gene copy number of X corresponds to $\log_{10}(X+1)$.

# CHAPTER 5. METAGENOMIC CHARACTERIZATION OF THE HOT WATER PLUMBING MICROBIOME AND RELATIONSHIP WITH OPPORTUNISTIC PATHOGEN OCCURRENCE

Pan Ji, William J. Rhoads, Marc A. Edwards, Amy Pruden

## ABSTRACT

Hot water plumbing delivers microbes into the built environment and represents a crucial element of public exposure to opportunistic pathogens (OPs). However, little is known about hot water plumbing microbiome, especially in terms of its functional composition. Two parallel lab-scale hot water rigs were constructed to explore hot water plumbing microbiome under different operation conditions and engineering control measures. Shotgun metagenomics was applied to a cross section of 36 select samples, allowing comparison between/among phases (biofilm and water), locations (influent, recirculation line and distal taps), water heater temperature set points (40 and 60 °C), water use frequencies (high- and low-, 21 and 1 flush per week, respectively), and three engineering control measures (heat shock, dosage of copper as antimicrobial, and dosage of chloramine). All samples shared highly similar taxonomic composition at the phylum level and functional genes profile at highest level (SEED Subsystem). Nonmetric Multidimensional Scaling ordination indicated distinct microbiomes in both taxonomic composition and functional gene profiles across locations. The Mantel test revealed a strong correlation between functional diversity and taxonomic diversity (Mantel $r = 0.602$, $P = 0.001$). Among the three engineering control measures tested, the mild heat shock conferred the least change in either the functional or taxonomic composition of the microbiome. Dosage of copper appeared to incur the most dramatic effect on the bulk water microbiome composition, while dosage of chloramine had more of an effect on biofilm than water. Positive associations were noted between all genera pairs including two amoeba genera and three OP containing genera.

## INTRODUCTION

Hot water plumbing is an important element of the building plumbing system, yet it differs from cold water plumbing in several characteristics in relevance to the microbiome. Hot water plumbing typically experiences high temperature, e.g., water heater thermostats mostly set between 49 and 60 °C in the U.S. (Department of Energy). In reality, hot water system is subject to temperature variation due to heat loss during flow and at-the-tap stagnation. A recent study reported temperature loss of up to 18.4 °C in the horizontal recirculation loop of a hot water system in a 400-bed university hospital in Sherbrooke, Canada (Bédard et al., 2016). In natural water environments, temperature has been recognized as a key factor which shapes microbial community (surface and deep water in the Gulf of Mexico, Redmond and Valentine, 2012; hot springs of the Tibetan Plateau, Wang et al., 2013b). In theory, the temperature variation within hot water system would resulted in diverse 'local' microbial communities. However, most efforts on deciphering the black box have been directed towards the microbial composition at point of use shower heads (Feazel et al., 2009) and shower hoses (Proctor et al., 2016), calling for a mechanistic understanding of the entire hot water plumbing microbiome.

There is growing interest in engineering hot water plumbing microbiome (Wang et al., 2013a; Dai et al., 2017), partly due to the escalating concerns on opportunistic pathogens (OPs) including *Legionella pneumophila* (Falkinham III et al., 2015; Schwake et al., 2016; Rhoads et al., 2017b), who are particularly challenging to control as they can colonize and proliferate within biofilms (Rogers et al., 1994; Declerck, 2009). Engineering control measures that have been adopted include heat shock or thermal disinfection (Bédard et al., 2016), copper-silver ionization (Stout and Yu, 2003), and on-site secondary disinfection (e.g., chloramine, Baron et al., 2015). Besides their efficacy on *L. pneumophila* control, little is known about the impact of these engineering control measures on hot water plumbing microbiome.

As the third part in this "trilogy" focused on the hot water microbiome (Chapters 3, 4, and 5), here we employed shotgun metagenomics seeking to understand the functional gene profile in addition to taxonomic composition of the hot water plumbing microbiome. A lab-scale recirculating hot water plumbing rig was constructed to isolate the impact of water heater

temperature set point and water use frequency. The parallel rig design allowed us to assess the impact of three engineering control measures: heat shock, dosage of copper as antimicrobial, and dosage of chloramine. With 36 selected samples, we aim to answer four questions in this preliminary data analysis: (1) what are the taxonomic compositions and functional gene profiles of hot water plumbing microbiome? (2) to what extent does hot water plumbing microbiome vary across locations? (3) can the three engineering control measures effectively mediate microbiome? (4) are there certain microbes closely associated with OPs?

## MATERIALS AND METHODS

**System description.** Two parallel lab-scale hot water rigs were installed to study hot water plumbing microbiome and its response to various engineering conditions. Both rigs were fed with Blacksburg, VA tap water passing through three consecutive granulated activated carbon (GAC, Pentek 10 inch Big Blue Model) filters to remove chloramine as disinfectant residual. Each rig consists of a 71.9 L water heater with recirculating loop, and 18 distal pipes in 2 pipe orientations (30° slanted from vertically upward vs vertically downward) and 3 water use frequencies (high, medium, low, as 21, 3, 1 flush/week) in triplicate. Three engineering control measures were tested: heat shock (water heater temperature set point at 60 °C and each set of distal pipes being intermittently flushed to achieve > 55 °C for 30 minutes), dosage of copper as antimicrobial, and dosage of chloramine. To commence this study (T0), both rigs had been acclimated for 4 months. Rigs #1 and #2 were maintained at 40 and 60 °C, respectively, for another 2 months (T1); then subject to the same heat shock procedure and both subsequently maintained at 40 °C for 2 months (T2). In the following 2 months, Rig #1 was dosed with cupric ion in the form of $CuSO_4$ targeting at a final concentration of 0.3 mg/L $Cu^{2+}$ (Cu), while Rig #2 remained undisturbed (T3). In the final stage, Rig #2 was dosed with chloramine at 1.5 mg/L (as total chlorine) for 2 months (Ch). At each of the five conditions (T0, T1, T2, T3, Cu and Ch), both first-flush bulk water (~500 mL) and biofilm (~ 65 $cm^2$ swabbed) samples were collected from every sampling portal in Rigs #1 and #2. Detailed experiment timeline is in Fig. 5.1.

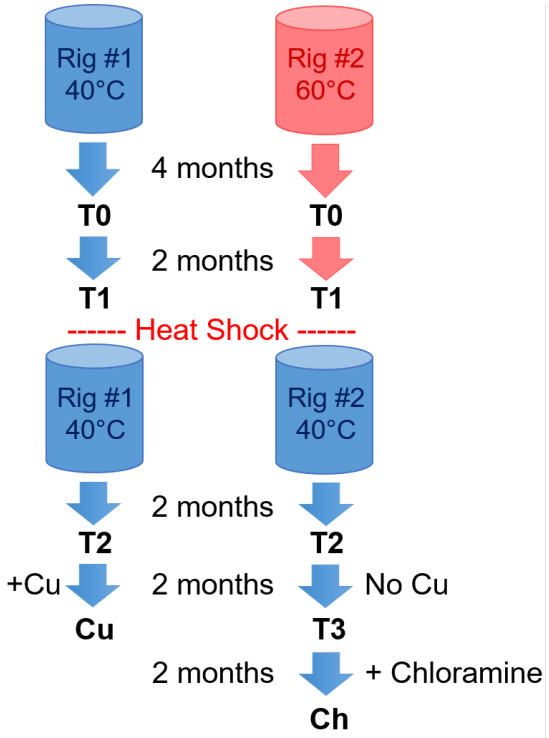


**Fig. 5.1. Experiment timeline.**

**Shotgun metagenomics and predicted metagenomics.** Water samples were first filter concentrated using 0.22 µm pore-size sterile mixed cellulose ester filters (Millipore, Billerica, MA, USA). All samples were then subject to DNA extraction using FastDNA™ SPIN kit (FastDNA™, MP Biomedicals, Solon, OH, USA). A total of 36 samples were selected for shotgun metagenomics, including 2 water samples collected at the influent portal from T0 (first-flush and flushed), 18 paired biofilm and water samples collected at recirculation line of both rigs, and 16 paired high and low water use frequency water samples from downward oriented pipes at T0, T1 and T2 (including 4 biological duplicates at T2). Shotgun metagenomics using Accel-NGS® 2S Plus DNA Library Kit (Swift Biosciences™, Ann Arbor, MI, USA) was performed on Illumina HiSeq 2500 platform (2 × 100 bp) at the Biocomplexity Institute at Virginia Tech. All distal tap samples tested were from downward oriented pipes, which in theory provide little convective mixing thus result in drastic contrast between high and low water use frequencies. Raw sequence reads were analyzed in MG-RAST (version 4.0.2), with sequencing depth of 21 516 114 ~ 40 374 025 per sample. Taxonomic hits distribution was obtained using protein similarities against the REFseq protein database (e-value = $10^{-5}$, %-identity 60, and minimum length 15), while functional category hits distribution was achieved

via SEED Subsystems hierarchies. If not specified, representative hit strategy was applied in data analysis with normalization on hits numbers.

**Data analysis.** Bray-Curtis distance matrix was calculated based on genus-level taxonomic table and Level 3 functional gene profile (package 'vegan' version 2.4-3, Oksanen et al., 2017) in R (version 3.4.1, R Core Team, 2017). Mantel test was applied as a measure of the similarity between two distance matrices (package 'vegan' version 2.4-3, Oksanen et al., 2017). Significant differences among sample categories in both taxonomic and functional profiles were determined through Permutational Multivariate Analysis of Variance (Adonis, Anderson 2001, permutation = 999, {vegan} version 2.3-0, Oksanen et al., 2016). Confounding dispersion effect within each sample category was accounted for using a multivariate homogeneity of groups dispersions (variances) analysis (betadisper, Anderson, 2006; package 'vegan' version 2.4-3, Oksanen et al., 2017). Non-metric Multidimensional Scaling (NMDS) was applied to visualize previously obtained distance matrices (package 'ggplot2', version 2.2.1, Wickham, 2009). Maximal information coefficient (MIC, package 'minerva' version 1.4.7, Davide et al., 2012) was calculated to estimate the strength of potential association among classified microbial genera, with Pearson correlation coefficient ('stats' version 3.4.1, R Core Team, 2017) calculated to determine whether the association is positive or negative. Histogram of MIC was constructed in package 'lattice' (version 0.20-35, Sarkar, 2008). Correlation graphs were constructed through package 'corrplot' (version 0.77, Wei and Simko, 2016).

## RESULTS

**Taxonomy composition of hot water plumbing microbiome.** In each sample, 98.5 – 99.2% of the total representative hits belonged to the Bacteria domain (10 750 075 to 22 411 419). Within Bacteria, the top 10 abundant phyla of all samples combined accounted for 94.8 – 99.2% of the bacterial representative hits in each sample, including: *Proteobacteria*, *Deinococcus-Thermus*, *Actinobacteria*, *Firmicutes*, *Bacteroidetes*, *Planctomycetes*, *Acidobacteria*, *Chloroflexi*, *Nitrospirae*, and *Cyanobacteria*. Two influent samples showed apparent variation in taxonomic composition at phylum level, mainly involving *Deinococcus-Thermus* relative abundance (Fig. 5.2). Paired biofilm and water samples typically shared the same abundant phyla yet differ in

relative abundance (recirculation line samples, e.g., Rig #1 at T0; Rig #2 at T0, Fig. 5.2). Recirculation line and distal tap water samples from the same rig shared more similar phylum composition at T0 and T1, when Rig #1 was at 40 °C and Rig #2 at 60 °C. Such pattern was not apparent at T2, when both rigs were maintained at 40 °C. Dosage of copper and chloramine were both more effective in altering phylum level taxonomic composition than heat shock.

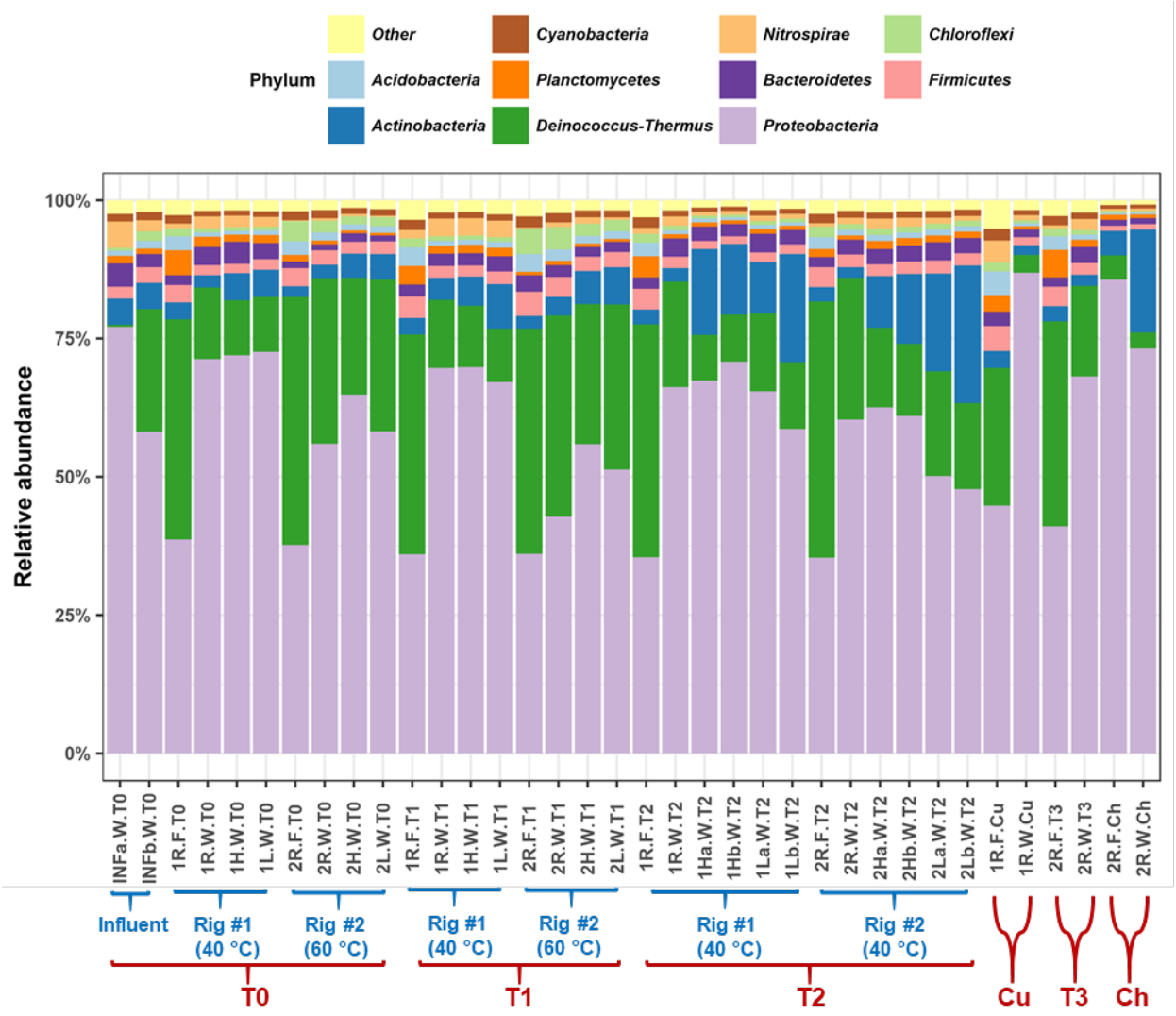


**Fig. 5.2. Top 10 abundant bacteria phyla of all samples combined.** Relative abundance was calculated as the ratio of representative hits number (within Bacteria domain). The rest phyla were combined as "Other".

At genus level, the top 10 abundant bacterial genera accounted for 34.1 – 56.1% of total representative hits in each sample. Among all samples, 41 out of 597 bacterial genera resembled

the top 10 abundant genera in individual samples (Fig. 5.3), 31 of which belong to phylum *Proteobacteria*. At T0 and T1, *Nitrospira* was a top abundant genus in influent and Rig #1 (40 °C), but not Rig #2 (60 °C), likely suggesting a preference for mild temperature. In biofilm phase at recirculation line in Rig #1, *Planctomyces* was consistently among the top abundant genera list (T0, T1, T2 and Cu). At T0 and T1, *Thermus* was more abundant in Rig #2 (60 °C) than in Rig #1 (40 °C) whereas *Meiothermus* showed the opposite trend. *Mycobacterium* was a top abundant genus frequently at influent and distal tap (T0, T1 and T2), and occasionally at recirculation line (only when chloramine was dosed, Ch).

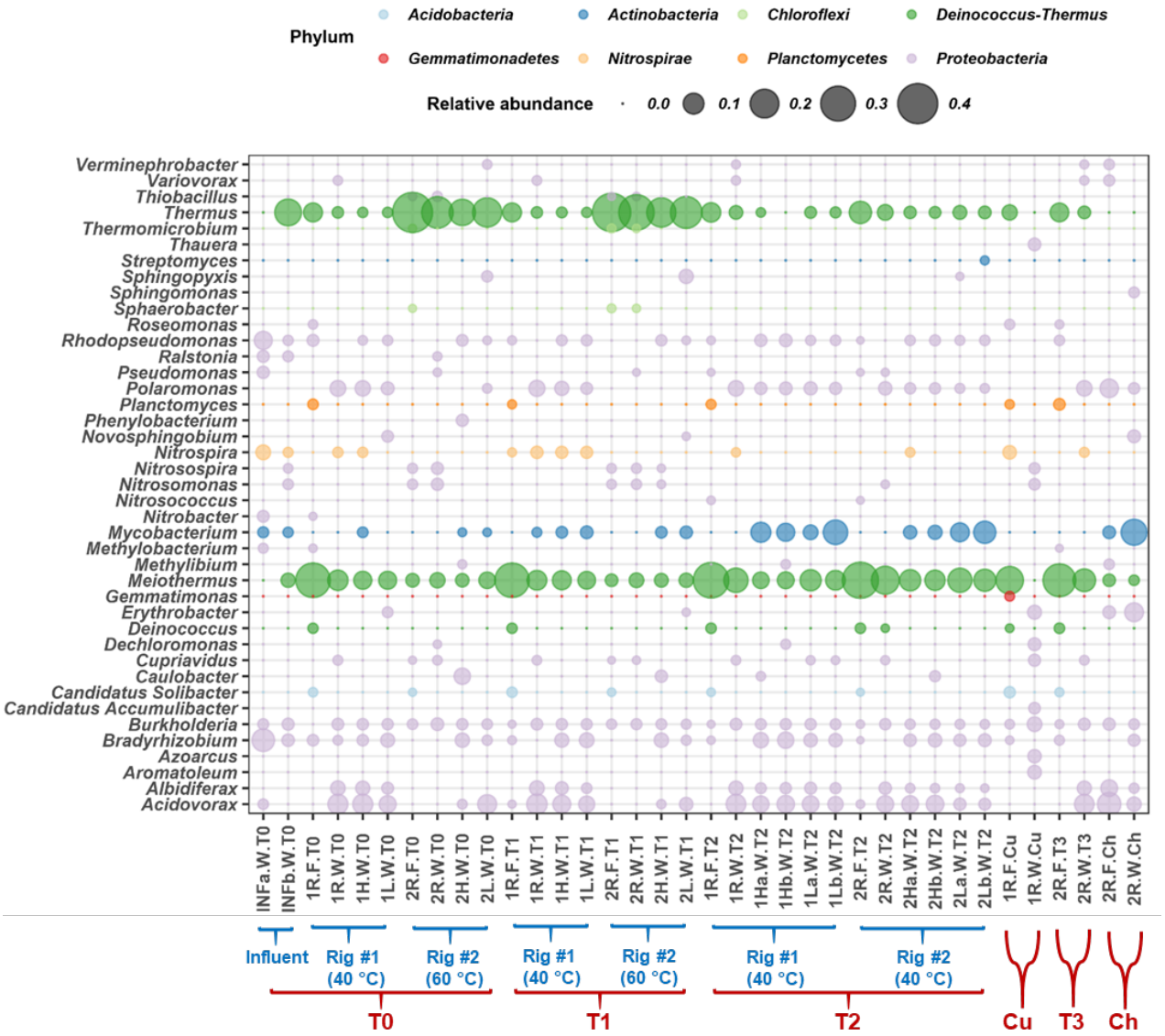


**Fig. 5.3. Top 10 abundant bacteria genera within each sample.** Point size indicates relative abundance and color coded by phylum.

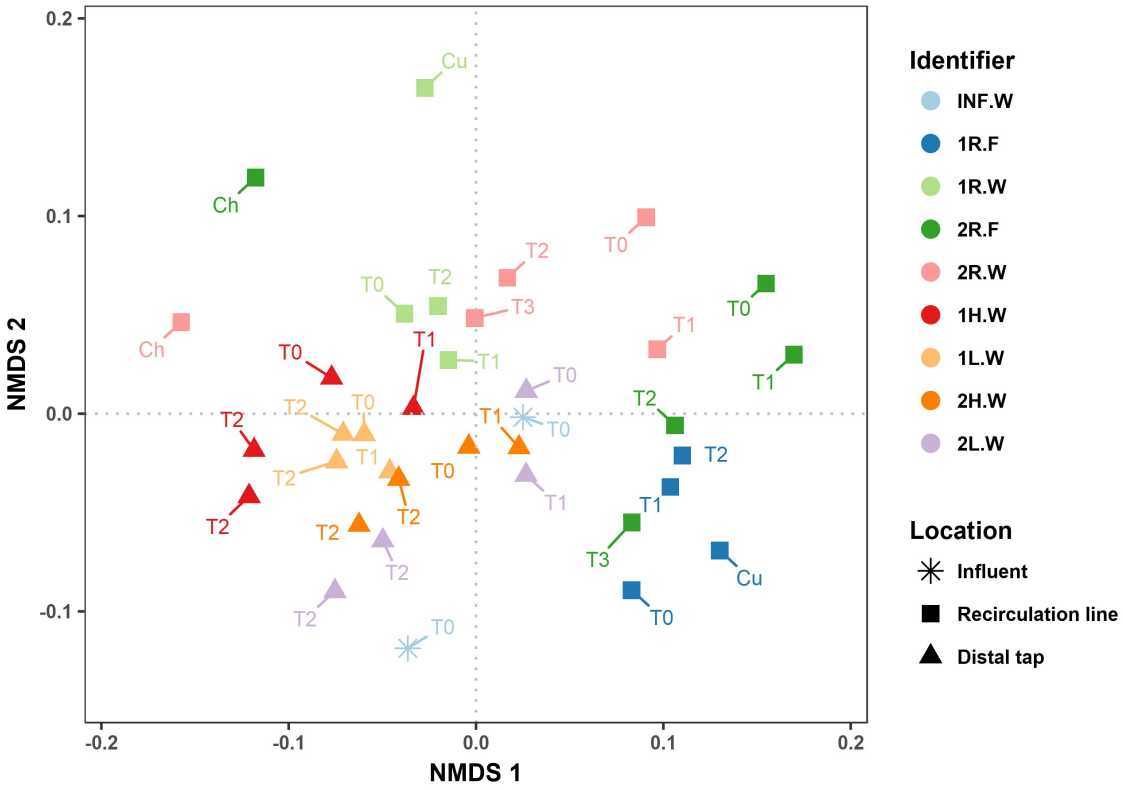


**Fig. 5.4. Nonmetric Multidimensional Scaling ordination plot based on genus level taxonomic composition (stress = 0.123).** Bray-Curtis distance matrix was used. Points (samples) shape coded by location and color coded by identifier, with condition labeled as text. Identifier aggregates information in rig (1 and 2), location (INF for influent, R for recirculation line, distal tap samples are denoted by water use frequency), water use frequency (H for high and L for low) and phase (F for biofilm and W for water).

Nonmetric multidimensional scaling (NMDS) ordination revealed that distal tap samples were clustered separately from recirculation line ones using Bray-Curtis distance matrix based on taxonomic composition at genus level (Fig. 5.4). At recirculation line, dosage of chloramine drastically changed bacterial community composition in both biofilm and bulk water phases (Ch; 2R.F and 2R.W), while dosage of copper as cupric ion was more effective in water than biofilm phase (Fig. 5.4). At distal tap, water microbial community remained distinct between rigs ($R^2_{Adonis}$ = 0.287, $P_{Adonis}$ = 0.001; $P_{Betadisper}$ = 0.001) and during heat shock process (T0, T1 and T2; $R^2_{Adonis}$ = 0.319, $P_{Adonis}$ = 0.022; $P_{Betadisper}$ = 0.134), while water use frequency exhibited no

significant impact ($R^2_{Adonis}$ = 0.075, $P_{Adonis}$ = 0.342; $P_{Betadisper}$ = 0.863). All four groups of biological duplicates at distal tap (T2; identifier 1H.W, 1L.W, 2H.W, and 2L.W) shared more similar microbial compositions than with non-biological duplicate samples from the same location and phase (Fig. 5.4). Notably, the difference in water microbiome at influent can be greater than that between different locations (distal tap vs recirculation line).

**Functional profile of hot water plumbing microbiome.** All samples shared similar major categories of functional genes (Level 1 by SEED Subsystems database, Fig. D.1). Statistically significant correlation (Mantel r = 0.602, P = 0.001) was found between the Bray-Curtis distance matrices calculated from taxonomy composition (representative hits at genus level) and functional gene abundances (Level 3 by SEED Sybsystems).

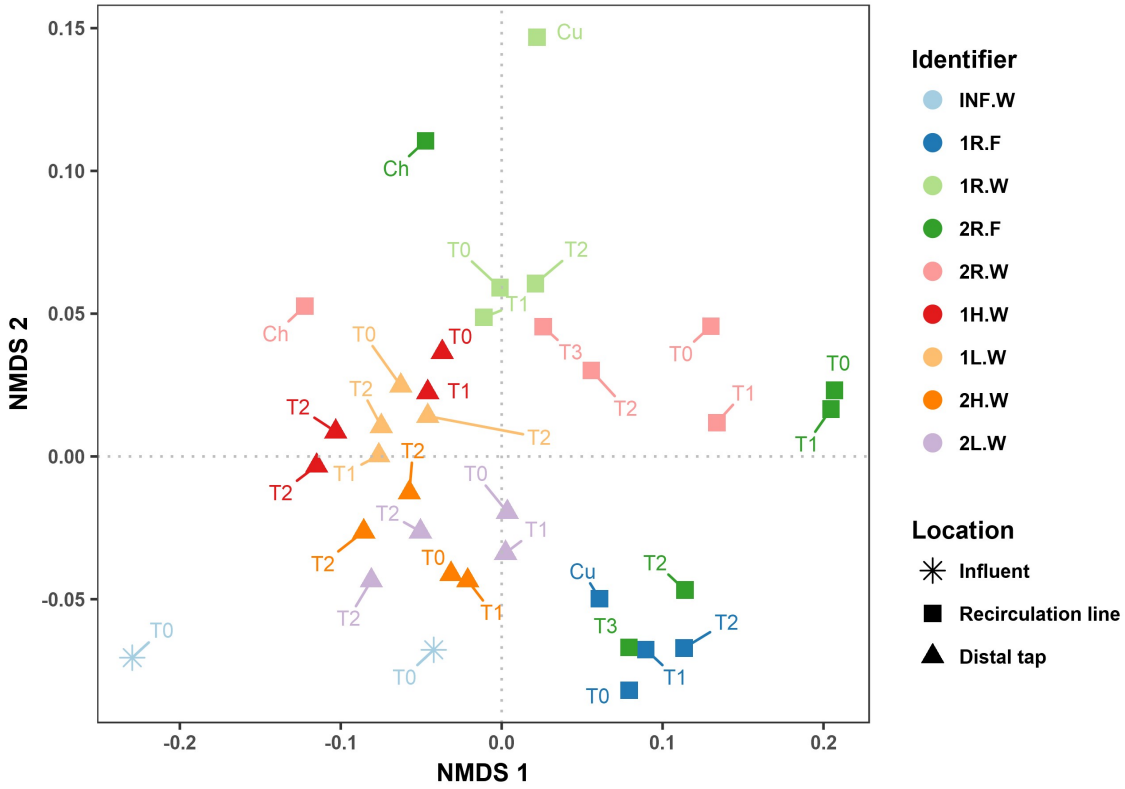


**Fig. 5.5. Nonmetric Multidimensional Scaling ordination plot based on Level 3 functional gene profile (stress = 0.110).** Bray-Curtis distance matrix was used. Points (samples) shape coded by location and color coded by identifier, with condition labeled as text. Identifier aggregates information in rig (1 and 2), location (INF for influent, R for recirculation line, distal

tap samples are denoted by water use frequency), water use frequency (H for high and L for low) and phase (F for biofilm and W for water).

Heat shock had limited impact on microbial functional profiles at recirculation line (Fig. 5.5, 1R.F, 1R.W; T0 and T1 vs T2), where decreased water heater temperature set point (60 to 40 °C) was more influential (2R.F, 2R.W; T0 and T1 vs T2). At the recirculation line, dosage of copper as cupric ion and dosage of chloramine both shifted microbial functional profile in biofilm and water phases, with greater change observed in water for the former and in biofilm for the latter. At the distal tap, water functional profile was distinct between the two rigs ($R^2_{Adonis}$ = 0.251, $P_{Adonis}$ = 0.014; $P_{Betadisper}$ = 0.181), yet not statistically different when comparing high to low water use frequency ($R^2_{Adonis}$ = 0.098, $P_{Adonis}$ = 0.223; $P_{Betadisper}$ = 0.649) or across three time points (T0, T1 and T2; $R^2_{Adonis}$ = 0.290, $P_{Adonis}$ = 0.058; $P_{Betadisper}$ = 0.913). The two influent water samples varied in functional gene profiles as well. During heat shock process, influent (water only), recirculation line (paired water and biofilm), and distal tap (water only) harbored statistically different functional gene profile ($R^2_{Adonis}$ = 0.296, $P_{Adonis}$ = 0.005; $P_{Betadisper}$ = 0.256).

**Associations between OP containing genera and other bacterial genera.** Maximal Information Coefficient (MIC) scores were calculated for each pair of the 563 classified bacterial genera (i.e. a total of 158 203 pairs), including 50 079 negative associations based on Pearson Correlation Coefficient (Fig. D.2, panel a). Both strong positive (e.g., *Nitrosospira* and *Nitrosomonas*, MIC = 0.910) and negative (e.g., *Meiothermus* and *Thermus*, MIC = -1.000) associations were observed among top 10 abundant bacterial genera for individual samples (Fig. 5.6). Of special interest are the associations involving 3 OP containing genera (i.e., *Legionella*, *Mycobacterium* and *Pseudomonas*), each group of which varied in proportion of negative association with other classified bacterial genera (17.2%, 50.7% and 34.4% negative associations, respectively; Fig. D.2, panels b – d). The top 20 positive and negative associations (ranked by MIC scores) involving the 3 OP containing genera were provided in Table 5.1. Interestingly, *Legionella* hits showed highest positive association with hits from two plant pathogen containing genera (*Xanthomonas*, MIC = 0.910; *Xylella*, MIC = 0.868) and human gastrointestinal tract bacteria containing genera (*Victivallis*, MIC = 0.820). In addition, strong

positive association was observed between *Legionella* hits and hits from two thermophilic bacterial genera, *Thermus* and *Chloroflexus* (MIC = - 0.732 for both).

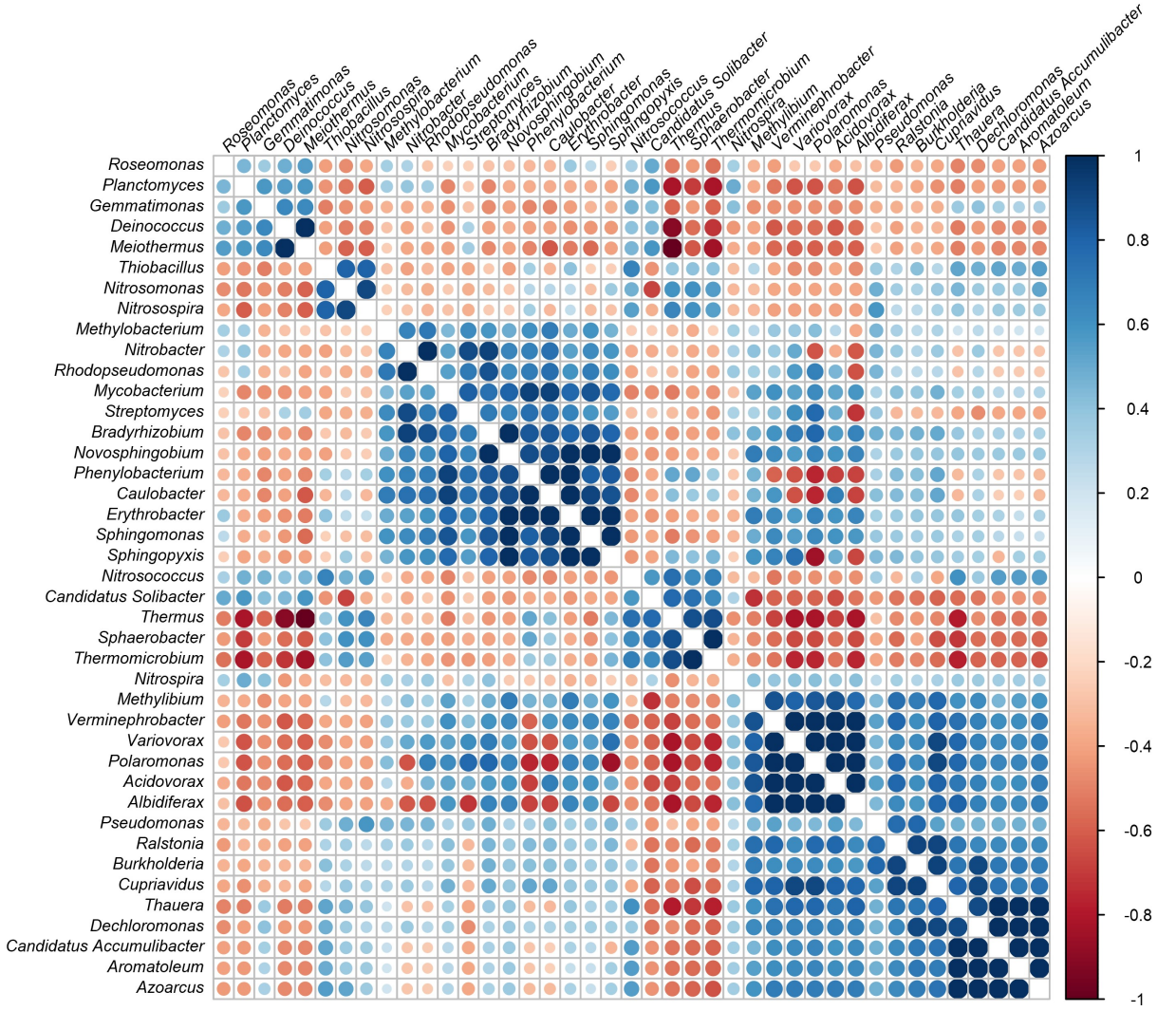


**Fig. 5.6. Heatmap of association among all top 10 bacteria genera for each sample.** Association strength was calculated as Maximal Information Coefficient (MIC) using relative abundance by representative hits numbers. Sign of association (positive or negative) was determined by Pearson Correlation Coefficient.

**Table 5.1. Top 20 positive and negative associations involving OP containing genera.** Ranking is based on Maximal Information Coefficient scores.

| | ***Legionella*** | | | | ***Mycobacterium*** | | | | ***Pseudomonas*** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Positive** | | **Negative** | | **Positive** | | **Negative** | | **Positive** | | **Negative** | |
| **Rank** | **Genus** | **MIC** | **Genus** | **MIC** | **Genus** | **MIC** | **Genus** | **MIC** | **Genus** | **MIC** | **Genus** | **MIC** |
| **1** | *Xanthomonas* | 0.910 | *Acidobacterium* | -0.732 | *Caulobacter* | 0.923 | *Oceanithermus* | -0.711 | *Yersinia* | 1.000 | *Leifsonia* | -0.484 |
| **2** | *Xylella* | 0.868 | *Candidatus Koribacter* | -0.732 | *Phenylobacterium* | 0.923 | *Pelobacter* | -0.658 | *Aeromonas* | 0.879 | *Olsenella* | -0.473 |
| **3** | *Victivallis* | 0.820 | *Chloroflexus* | -0.732 | *Asticcacaulis* | 0.910 | *Pelotomaculum* | -0.567 | *Cronobacter* | 0.789 | *Pediococcus* | -0.459 |
| **4** | *Moritella* | 0.791 | *Thermus* | -0.732 | *Gordonia* | 0.870 | *Acetohalobium* | -0.549 | *Burkholderia* | 0.783 | *Ktedonobacter* | -0.459 |
| **5** | *Stenotrophomonas* | 0.771 | *Dehalococcoides* | -0.641 | *Tsukamurella* | 0.870 | *Jonquetella* | -0.540 | *Escherichia* | 0.771 | *Anoxybacillus* | -0.459 |
| **6** | *Wolbachia* | 0.747 | *Terriglobus* | -0.641 | *Sagittula* | 0.870 | *Gallionella* | -0.528 | *Ralstonia* | 0.770 | *Thermoanaerobacter* | -0.459 |
| **7** | *Reinekea* | 0.730 | *Thermomicrobium* | -0.639 | *Brevundimonas* | 0.865 | *Syntrophobacter* | -0.527 | *Xenorhabdus* | 0.724 | *Syntrophothermus* | -0.459 |
| **8** | *Albidiferax* | 0.730 | *Candidatus Desulforudis* | -0.638 | *Sphingobium* | 0.856 | *Salinibacter* | -0.520 | *Salmonella* | 0.699 | *Caldanaerobacter* | -0.459 |
| **9** | *Aggregatibacter* | 0.727 | *Oceanithermus* | -0.625 | *Sphingomonas* | 0.856 | *Candidatus Desulforudis* | -0.519 | *Dickeya* | 0.673 | *Thermotoga* | -0.459 |
| **10** | *Magnetococcus* | 0.727 | *Thermobaculum* | -0.620 | *Zymomonas* | 0.856 | *Sideroxydans* | -0.518 | *Shewanella* | 0.673 | *Carboxydothermus* | -0.459 |
| **11** | *Planctomyces* | 0.708 | *Ammonifex* | -0.602 | *Dermacoccus* | 0.848 | *Alkalilimnicola* | -0.518 | *Citrobacter* | 0.673 | *Herpetosiphon* | -0.459 |
| **12** | *Chryseobacterium* | 0.700 | *Dictyoglomus* | -0.579 | *Kocuria* | 0.848 | *Thermus* | -0.514 | *Psychrobacter* | 0.668 | *Symbiobacterium* | -0.459 |
| **13** | *Photorhabdus* | 0.700 | *Ktedonobacter* | -0.579 | *Amycolatopsis* | 0.820 | *Collimonas* | -0.513 | *Vibrio* | 0.668 | *Thermaerobacter* | -0.459 |
| **14** | *Psychromonas* | 0.697 | *Moorella* | -0.579 | *Saccharomonospora* | 0.820 | *Hydrogenobacter* | -0.510 | *Photorhabdus* | 0.665 | *Blastopirellula* | -0.448 |
| **15** | *Enhydrobacter* | 0.696 | *Pediococcus* | -0.579 | *Segniliparus* | 0.811 | *Nitratiruptor* | -0.508 | *Providencia* | 0.661 | *Kineococcus* | -0.443 |
| **16** | *Basfia* | 0.689 | *Methylobacillus* | -0.573 | *Intrasporangium* | 0.811 | *Blastopirellula* | -0.502 | *Cellvibrio* | 0.658 | *Actinomyces* | -0.442 |
| **17** | *Glaciecola* | 0.689 | *Pelotomaculum* | -0.564 | *Nocardia* | 0.811 | *Pirellula* | -0.502 | *Achromobacter* | 0.657 | *Hydrogenobacter* | -0.441 |
| **18** | *Algoriphagus* | 0.688 | *Thiobacillus* | -0.561 | *Nocardioides* | 0.811 | *Chloroflexus* | -0.501 | *Pectobacterium* | 0.651 | *Candidatus Solibacter* | -0.432 |
| **19** | *Gramella* | 0.688 | *Streptobacillus* | -0.554 | *Frankia* | 0.811 | *Anaerobaculum* | -0.500 | *Janthinobacterium* | 0.651 | *Nostoc* | -0.426 |
| **20** | *Zunongwangia* | 0.688 | *Sideroxydans* | -0.535 | *Catenulispora* | 0.811 | *Nitrosococcus* | -0.499 | *Afipia* | 0.636 | *Spirochaeta* | -0.426 |

Representative hits matching genera *Acanthamoeba* and *Hartmannella* (including *Vermamoeba vermiformis*) were detected in 36/36 and 34/36 samples, both positively associated with all three OP containing genera. Compared to *Acanthamoeba*, *Hartmannella* showed stronger association with OP containing genera by representative hits, with the strongest association occurred with *Legionella* (MIC = 0.628; Fig. 5.7). The three OP containing genera were positively associated with each other as well. Of the 560 non-OPs containing classified bacterial genera and 2 amoeba genera (*Acanthamoeba* and *Hartmannella*), 234 were only positive associated with the 3 OP containing genera (Table D.1), and 68 were only negative associated with the 3 OP containing genera (Table D.2), with the rest 260 having mixed associations with the 3 OP containing genera. Among the 68 bacterial genera negatively associated with all 3 OP containing genera, none had MIC scores all above 0.50, yet 45 had MIC scores all above 0.30. For the 234 genera positively associated with all 3 OP containing genera, 4 had MIC scores all above 0.50 (*Acidovorax*, *Ketogulonicigenium*, *Stenotrophomonas* and *Verminephrobacter*), and 161 had MIC scores all above 0.30.

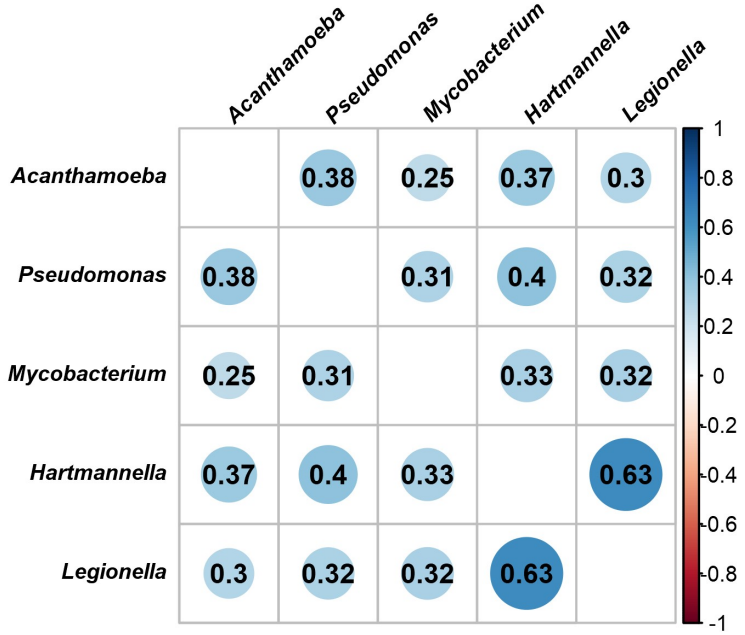


**Fig. 5.7. Heatmap of association among 2 amoeba and 3 OP containing genera. Association strength was calculated as Maximal Information Coefficient (MIC) using relative abundance by representative hits numbers.** Sign of association (positive or negative) was determined by Pearson Correlation Coefficient.

## DISCUSSION

**Hot water plumbing microbiome.** This study explores plumbing microbiome in a lab-scale recirculating hot water system that experienced three engineering control measures sequentially: heat shock, dosage of copper as antimicrobial, and dosage of chloramine. To the authors' knowledge, this is the first attempt to probe hot water plumbing microbiome functional gene profile via shotgun metagenomics. All 36 samples selected, representing both biofilm and water phases, different locations (influent, recirculation line, distal tap), and conditions under three engineering control processes, differed with each other mostly in terms of relative abundance (calculated by representative hits) rather than membership for top 10 abundant bacteria phyla (Fig. 5.2) and shared highly similar major functional gene categories (Level 1, Fig. D.1). A prior study conducted on the same system (Ji et al., 2017b) identified similar top 10 abundant bacteria phyla through Illumina 16S rRNA gene amplicon sequencing.

Greater variations were observed at genus level for taxonomic composition (Fig. 5.3). Different locations in the system (influent, recirculation line and distal tap) favor different top abundant genus, as is the case in *Mycobacterium*. NMDS ordination plots (Fig. 5.4 and Fig. 5.5) indicated that bacteria community varied across locations in terms of both genus-level taxonomic and Level 3 functional gene compositions. The three locations distinct in many aspects, which mainly involve temperature and flow patterns. Influent is always characterized with low to room temperature and recirculation line is consistently at water heater temperature set point. Distal taps routinely hit water heater temperature set point during flushing, then subsequently cool down to room temperature within 25 minutes after flushing. In terms of flow pattern, recirculation line is subject to continuous flow and well mixed with water heater tank. Distal tap typically subject to standard flush/ stagnant cycle based on water use frequency. Importantly, in this study flow/ stagnant cycle occurred at influent concurrently with flushing at any set of distal taps, which likely explained the large difference between first-flush and flushed influent waters. Notably, such difference can be greater than that those between influent and distal tap.

**Effectiveness of the three engineering control measures on microbiome mediation.** Heat shock or thermal disinfection (Bédard et al., 2016), copper as antimicrobial (Stout and Yu, 2003;

Rhoads et al., 2017a) and chloramine (Baron et al., 2015) each has been considered or hypothesized as an effective measure for OPs control in building plumbing systems. Among the three, heat shock appeared to be the least effective in altering either taxonomic or functional gene composition at recirculation line, which corroborates prior Illumina 16S rRNA gene amplicon sequencing results (Ji et al., 2017b). It should be emphasized that the heat shock procedure selected was on the mild end of the available spectrum of published protocols (AWT, 2003; Florida DOH, 2014; OSHA, 2014). Consistent with previous findings (Ji et al., 2017b), water heater temperature set point (60 vs 40 °C) appeared to be more influential on microbiome composition (both taxonomy and functional genes) than heat shock treatment.

In this study, copper dosing as cupric ion mostly impacted microbiome in water phase, while chloramine dosage affected both biofilm and water phases with more drastic change observed in biofilm phase (Fig. 5.4 and Fig. 5.5). Despite being recognized as an antimicrobial agent, copper has been reported to be less effective towards eradication or inactivation of cells in biofilm than in water phase (Teitzel and Parsek, 2003; Harrison et al., 2005), putatively due to the protection provided through sorption and sequestration of copper ion by extracellular polymeric substance (EPS) components (Hunt, 1986; Chen et al., 2011). Monochloramine was found to outcompete chlorine in terms of biofilm penetration (Lee et al., 2011), which likely explained its effect on biofilm microbiome. Chiao et al. (2014) found that viable bacterial community showed greater changes than total bacterial community under monochloramine inactivation. This study focuses on the total bacterial community, and might underrepresent change in active function profile.

**Microbial genera with strong associations to OP containing genera.** A long-term goal for OP control is to establish a probiotic approach which mediates the entire microbiome in building plumbing system (Wang et al., 2013a). Here we applied MIC as a measure of association strength between OP containing genera and other bacterial genera using relative abundance calculated as ratio of representative hits. The sign of association was determined by Pearson Correlation Efficient. It needs be noted that the analysis results should not be considered as proof of microbial ecological relationship, but rather, indicators that certain genera might contribute to enhance or inhibit OP regrowth.

Positive associations were observed between 3 OP containing genera and 2 amoeba genera, with the strongest association occurred between *Legionella* and *Hartmannella* (MIC = 0.628). A field study reported low to moderate correlation between *Legionella* spp. and *Hartmannella vermiformis* (née *Vermamoeba vermiformis*) using quantitative PCR in one of the two chloraminated drinking water distribution systems (Wang et al., 2012). The 3 OP containing genera were also positively associated with each other, where positive association has been noted between *Legionella* and *Mycobacterium* in a prior study on the rig system (Ji et al., 2017a, via Illumina 16S rRNA gene amplicon sequencing). A question of special interest is if there exist bacterial genera positively or negatively associated with all 3 OP containing genera. Our results identified a total of 234 bacterial and amoeba genera who were only positive associated with the 3 OP containing genera, and 68 bacterial genera only negative associated with the 3 OP containing genera. Notably, in either scenario, no genus showed strong association with all 3 OP containing genera (MIC > 0.7), suggesting a microbial ecology more complex than single genus as OPs enhancer or inhibitor.

## CONCLUSIONS

This study is the first to our knowledge to attempt to explore the hot water plumbing microbiome with a shot-gun metagenomics approach. Through implementing lab-scale parallel hot water rigs, both taxonomic compositions and functional gene profiles were captured in both biofilm and water phases at influent, recirculation line and distal taps. Strong correlation was found between functional diversity and taxonomic diversity. Three engineering control measures were tested for their effectiveness on microbiome mediation. The mild heat shock appeared to be the least influential, with copper dosing mostly affecting water and chloramine dosage more impactful in biofilm than water. Positive association was observed between amoeba genera and OP containing genera, with other bacterial genera highly associated with OPs noted as well.

## REFERENCES

Anderson MJ. (2001). A new method for non-parametric multivariate analysis of variance. *Aust Ecol* 26: 32-46.

Anderson MJ. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* 62(1): 245–253.

Association of Water Technologies (AWT). (2003). Legionella 2003: An Update and Statement by the Association of Water Technologies. Association of Water Technologies. https://www.awt.org/pub/035C2942-03BE-3BFF-08C3-4C686FB7395C. Accessed 08/20/2017.

Baron JL, Harris JK, Holinger EP, Duda S, Stevens MJ, Robertson CE, et al. (2015). Effect of monochloramine treatment on the microbial ecology of *Legionella pneumophila* and associated bacterial populations in a hospital hot water system. *Syst Appl Microbiol* 38: 198 – 205. doi: 10.1016/j.syapm.2015.02.006.

Bédard E, Boppe I, Koumamé S, Martin P, Pinsonneault L, Valiquette L, et al. (2016). Combination of heat shock and enhanced thermal regime to control the growth of a persistent *Legionella pneumophila* strain. *Pathogens* 5(2): 35. doi:10.3390/pathogens5020035.

Chen G, Chen X, Yang Y, Hay AG, Yu X, Chen Y. (2011). Sorption and distribution of copper in unsaturated *Pseudomonas putida* CZ1 biofilms as determined by X-Ray Fluorescence microscopy. *Appl Environ Microbiol* 77(14): 4719 – 4747. doi: 10.1128/AEM.00125-11.

Chiao T-H, Clancy TM, Pinto A, Xi C, Raskin L. (2014). Differential resistance of drinking water bacterial populations to monochloramine disinfection. *Environ Sci Technol* 48(7): 4038 – 4047. doi: 10.1021/es4055725.

Dai D, Prussin AJ, Mar LC, Vikesland PJ, Edwards MA, Pruden A. (2017). Factors shaping the human exposome in the built environment: opportunities for engineering control. *Environ Sci Technol* 51: 7759 – 7774. doi: 10.1021/acs.est.7b01097.

Davide A, Filosi M, Visintainer R, Riccadonna S, Jurman G, and Furlanello C. (2013). Minerva and minepy: a C engine for the MINE suite and its R, Python and MATLAB wrappers. *Bioinformatics* 29(3): 407 – 408. doi: 10.1093/bioinformatics/bts707.

Declerck P. (2009). Biofilm: the environmental playground of *Legionella pneumophila*. *Environ Microbiol* 12(8): 557 – 566. doi: 10.1111/j.1462-2920.2009.0205.x.

Department of Energy. Savings project: lower water heating temperature. https://www.energy.gov/energysaver/projects/savings-project-lower-water-heating-temperature. Accessed 08/30/2017.

Earth Microbiome Project. http://www.earthmicrobiome.org/protocols-and-standards/16s/. Accessed 08/20/2017.

Falkinham III JO, Pruden A, Edwards M. (2015). Opportunistic premise plumbing pathogens: increasingly important pathogens in drinking water. *Pathogens* 4: 373 – 386. doi: 10.3390/pathogens4020373.

Feazel LM, Baumgartner LK, Peterson KL, Frank DN, Harris JK, Pace NR. (2009). Opportunistic pathogens enriched in showerhead biofilms. *PNAS* 106(38): 16393 – 16399. doi: 10.1073/pnas.0908446106.

Florida Department of Health (Florida DOH). (2014). Guidelines for the Surveillance, Investigation, and Control of Legionnaires' Disease in Florida. http://www.floridahealth.gov/diseases-and-conditions/legionnaires-disease/_documents/gsi-legionella-update-final.pdf. Accessed 05/22/2017.

Harrison JJ, Turner RJ, Ceri H. (2005). Persister cells, the biofilm matrix and tolerance to metal cations in biofilm and planktonic *Pseudomonas aeruginosa*. *Environ Microbiol* 7(7): 981 – 994. doi: 10.1111/j.1462-2920.2005.00777.x.

Hunt S. (1986). Diversity of biopolymer structure and its potential for ion-binding applications. 15-46. In Eccles H and Hunt S(ed.), Immobilization of ions by bio-sorption. Ellis Horwood Ltd., West Sussex, United Kingdom.

Ji P, Rhoads WJ, Edwards MA, Pruden A. (2017a). Impact of water heater temperature setting and water use frequency on the building plumbing microbiome. *ISME J* 11: 1318 – 1330. doi: 10.1038/ismej.2017.14.

Ji P, Rhoads WJ, Edwards MA, Pruden A. (2017b). Effect of heat shock on hot water plumbing microbiota and *Legionella pneumophila* control. *Microbiome*. In Revision.

Lee WH, Wahman DG, Bishop PL, Pressman JG. (2011). Free chlorine and monochloramine application to nitrifying biofilm: comparison of biofilm penetration, activity, and viability. *Environ Sci Technol* 45(4): 1412 – 1419. doi: 10.1021/es1035305.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. (2017). vegan: Community Ecology Package. R package version 2.4-3. https://CRAN.R-project.org/package=vegan. Accessed 08/20/2017.

Occupational Safety and Health Administration (OSHA). (2014). OSHA Technical Manual TED 01-00-015. https://www.osha.gov/dts/osta/otm/otm_iii/otm_iii_7.html#5. Accessed 05/22/2017.

Proctor C, Gächter M, Kötsch S, Rölli F, Sigrist R, Walser J-C, et al. (2016). Biofilms in shower hoses – choice of pipe material influences bacterial growth and communities. *Environ Sci: Water Res Technol* 2: 670 – 682. doi: 10.1039/C6EW00016A.

R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Redmond MC, Valentine DL. (2012). Natural gas and temperature structured a microbial community response to the *Deepwater Horizon* oil spill. *PNAS* 109(50): 20292 – 20297. doi: 10.1073/pnas.1108756108.

Rhoads WJ, Pruden A, Edwards MA. (2017a). Interactive effects of corrosion, copper, and chloramines on *Legionella* and *Mycobacteria* in hot water plumbing. *Environ Sci Technol* 51(12): 7065 – 7075. doi: 10.1021/acs.est.6b05616.

Rhoads WJ, Garner E, Ji P, Zhu N, Parks J, Schwake D, et al. (2017). Distribution system operational deficiencies coincide with reported Legionnaires' disease clusters in Flint, MI. *Environ Sci Technol*. Accepted.

Rogers J, Dowsett AB, Dennis PJ, Lee JV, Keevil CW. (1994). Influence of temperature and plumbing material selection on biofilm formation and growth of *Legionella pneumophila* in a model potable water system containing complex microbial flora. *Appl Environ Microbiol* 5: 1585 – 1592.

Sarkar D. (2008). Lattice: Multivariate Data Visualization with R. Springer, New York. ISBN 978-0-387-75968-5.

Schwake DO, Garner E, Strom OR, Pruden A, Edwards MA. (2016). *Legionella* DNA markers in tap water coincident with a spike in Legionnaires' disease in Flint, MI. *Environ Sci Technol Lett* 3(9): 311 – 315. doi: 10.1021/acs.estlett.6b00192.

Stout JE, Yu VL. (2003). Experiences of the first 16 hospitals using copper-silver ionization for *Legionella* control: implications for the evaluation of other disinfection modalities. *Infec Control Hosp Epidemiol* 24(8): 563 – 568.

Teitzel GM, Parsek MR. (2003). Heavy metal resistance of biofilm and planktonic *Pseudomonas aeruginosa*. *Appl Environ Microbiol* 69(4): 2313 – 2320. doi: 10.1128/AEM.69.4.2313-2320.2003.

Wang H, Edwards M, Falkinham III JO, Pruden A. (2012). Molecular survey of the occurrence of *Legionella* spp., *Mycobacterium* spp., *Pseudomonas aeruginosa*, and amoeba hosts in two chloraminated drinking water distribution system. *Appl Environ Microbiol* 78(17): 6285 – 6294. doi: 10.1128/AEM.01492-12.

Wang H, Edwards MA, Falkinham III JO, Pruden A. (2013a). Probiotic approach to pathogen control in premise plumbing systems? A review. *Environ Sci Technol* 47: 10117 – 10128. doi: 10.1021/es402455r.

Wang S, Hou W, Dong H, Jiang H, Huang L, Wu G, et al. (2013b). Control of temperature on microbial community structure in hot springs of the Tibetan Plateau. *PLoS One* 8(5): e62901. doi: 10.1371/journal.pone.0062901.

Wei T, and Simko V. (2016). corrplot: Visualization of a Correlation Matrix. R package version 0.77. https://CRAN.R-project.org/package=corrplot. Accessed 08/27/2017.

Wickham H. (2009). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

## APPENDIX D – SUPPLEMENTARY INFORMATION FOR CHAPTER 5

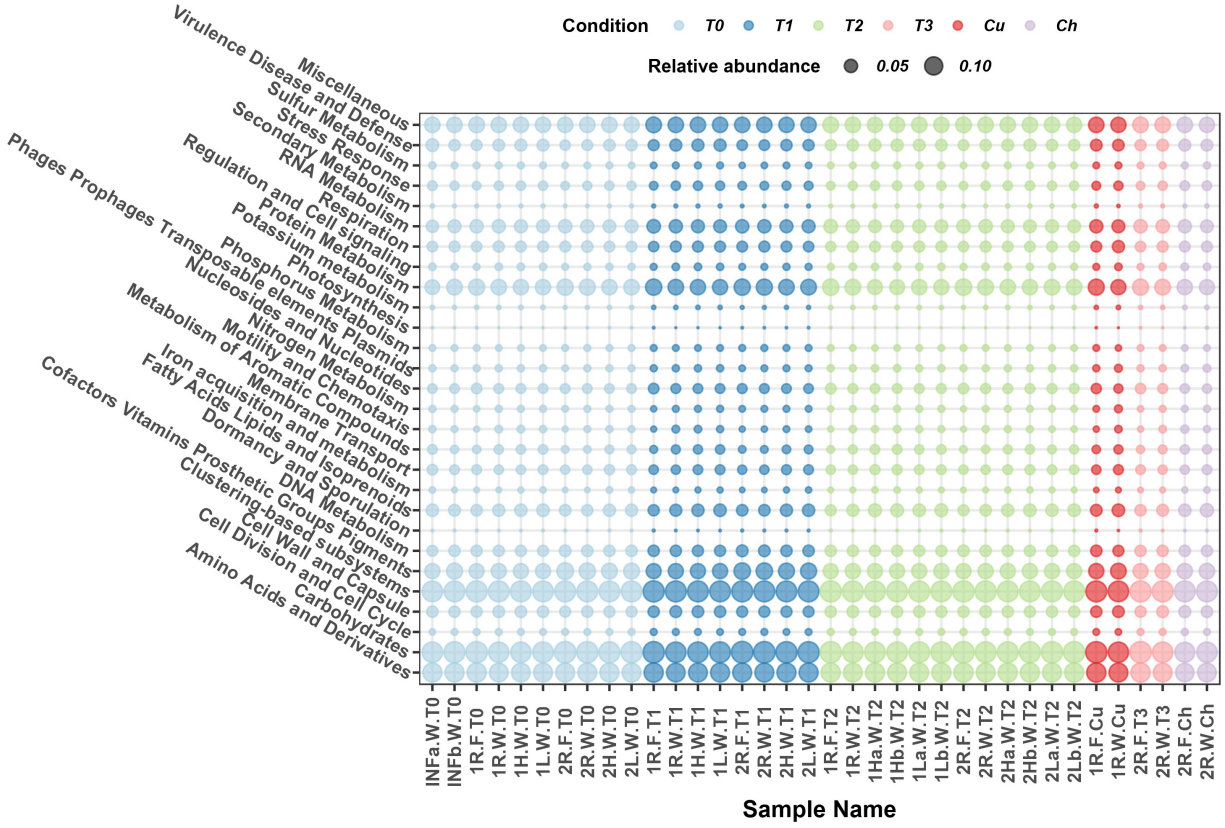


**Fig. D.1. Major functional groups across samples.** Level 1 functional gene profile was used. Point size indicates relative abundance by ratio of representative hits numbers, with points color coded by condition.

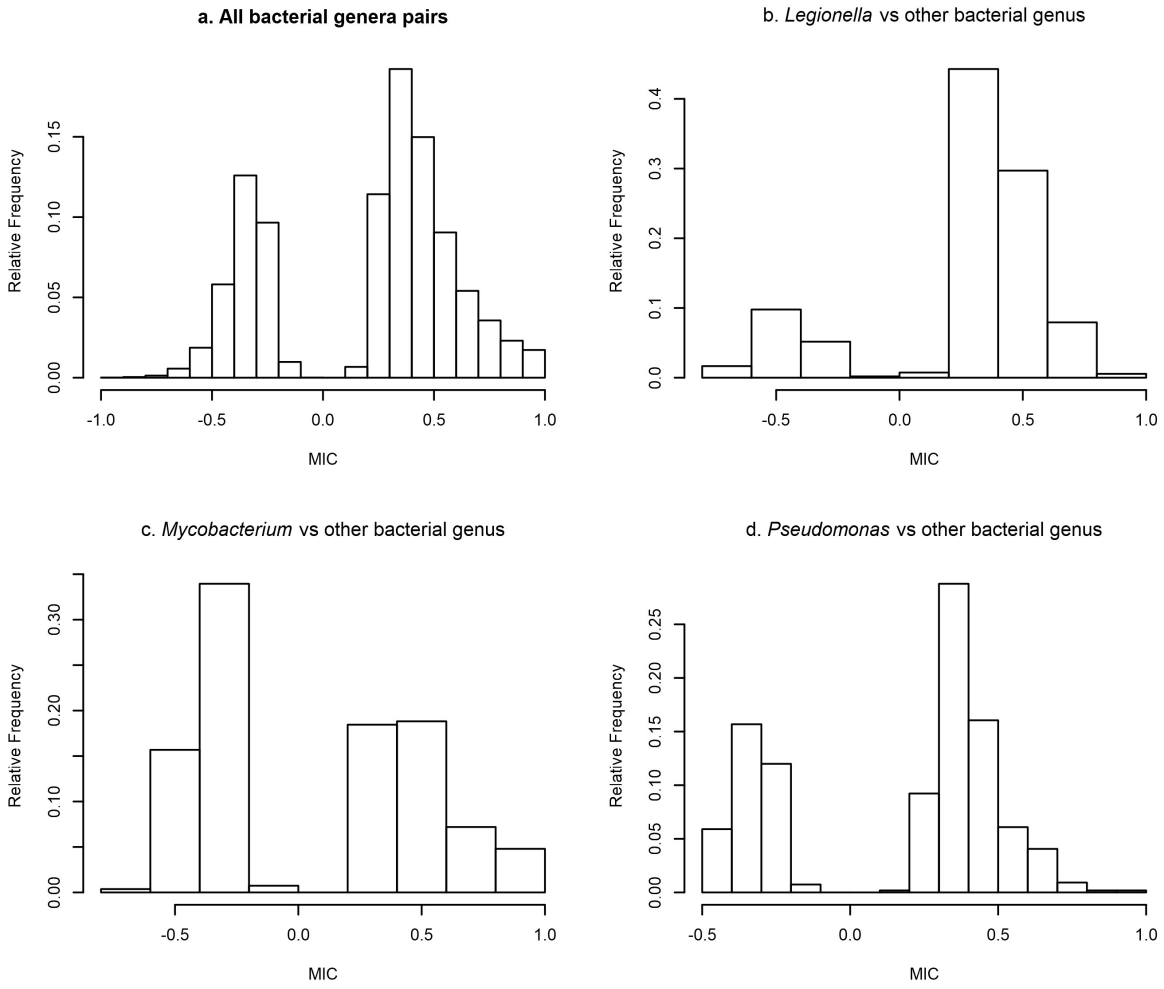


**Fig. D.2. Histogram on Maximal Information Coefficient (MIC) scores distributions.** Sign of association (positive or negative) was determined by Pearson Correlation Coefficient.

**Table D.1. Classified bacterial and 2 amoeba genera positively associated with 3 OP containing genera.** Association strength was calculated via Maximal Information Coefficient (MIC) using representative hits. Sign of association (positive or negative) was determined through Pearson Correlation Coefficient. Rank was based on MIC scores across all genera in the table. *Legionella*, *Mycobacterium* and *Pseudomonas* has been abbreviated as "*Leg.*", "*Myco.*" and "*Pseu.*", respectively.

| Genus | MIC score with | | | Ranked by MIC score | | | Domain | Phylum | Class | Order | Family |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Leg.* | *Myco.* | *Pseu.* | *Leg.* | *Myco.* | *Pseu.* | | | | | |
| *Acanthamoeba* | 0.296 | 0.254 | 0.375 | 198 | 223 | 151 | Eukaryota | unclassified (derived from Eukaryota) | unclassified (derived from Eukaryota) | unclassified (derived from Eukaryota) | unclassified (derived from Eukaryota) |
| *Acaryochloris* | 0.471 | 0.397 | 0.497 | 92 | 142 | 60 | Bacteria | Cyanobacteria | unclassified (derived from Cyanobacteria) | unclassified (derived from Cyanobacteria) | unclassified (derived from Cyanobacteria) |
| *Acholeplasma* | 0.520 | 0.299 | 0.301 | 69 | 201 | 219 | Bacteria | Tenericutes | Mollicutes | Acholeplasmatales | Acholeplasmataceae |
| *Achromobacter* | 0.279 | 0.384 | 0.657 | 210 | 150 | 17 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Alcaligenaceae |
| *Acidimicrobium* | 0.310 | 0.518 | 0.357 | 187 | 84 | 171 | Bacteria | Actinobacteria | Actinobacteria (class) | Acidimicrobiales | Acidimicrobiaceae |
| *Acidovorax* | 0.552 | 0.530 | 0.543 | 53 | 79 | 40 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Comamonadaceae |
| *Acinetobacter* | 0.548 | 0.410 | 0.573 | 56 | 132 | 36 | Bacteria | Proteobacteria | Gammaproteobacteria | Pseudomonadales | Moraxellaceae |
| *Actinosynnema* | 0.388 | 0.750 | 0.292 | 130 | 32 | 224 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Actinosynnemataceae |
| *Aeromicrobium* | 0.322 | 0.791 | 0.306 | 175 | 23 | 216 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Nocardioidaceae |
| *Aeromonas* | 0.369 | 0.306 | 0.879 | 144 | 199 | 2 | Bacteria | Proteobacteria | Gammaproteobacteria | Aeromonadales | Aeromonadaceae |
| *Afipia* | 0.246 | 0.518 | 0.636 | 230 | 83 | 20 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Bradyrhizobiaceae |
| *Aggregatibacter* | 0.727 | 0.371 | 0.518 | 8 | 156 | 55 | Bacteria | Proteobacteria | Gammaproteobacteria | Pasteurellales | Pasteurellaceae |
| *Agrobacterium* | 0.257 | 0.531 | 0.384 | 222 | 76 | 143 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiaceae |
| *Ahrensia* | 0.255 | 0.470 | 0.388 | 227 | 101 | 132 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Albidiferax* | 0.730 | 0.584 | 0.441 | 7 | 55 | 98 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Comamonadaceae |
| *Alcanivorax* | 0.459 | 0.232 | 0.471 | 103 | 227 | 79 | Bacteria | Proteobacteria | Gammaproteobacteria | Oceanospirillales | Alcanivoracaceae |
| *Algoriphagus* | 0.688 | 0.417 | 0.281 | 14 | 129 | 227 | Bacteria | Bacteroidetes | Cytophagia | Cytophagales | Cyclobacteriaceae |
| *Alicycliphilus* | 0.590 | 0.543 | 0.441 | 43 | 65 | 99 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Comamonadaceae |
| *Alistipes* | 0.661 | 0.477 | 0.339 | 25 | 97 | 187 | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Rikenellaceae |
| *Alteromonas* | 0.386 | 0.357 | 0.486 | 131 | 161 | 62 | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | Alteromonadaceae |
| *Amycolatopsis* | 0.339 | 0.820 | 0.306 | 164 | 9 | 211 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Pseudonocardiaceae |
| *Anaplasma* | 0.467 | 0.438 | 0.278 | 94 | 116 | 229 | Bacteria | Proteobacteria | Alphaproteobacteria | Rickettsiales | Anaplasmataceae |
| *Arcobacter* | 0.653 | 0.333 | 0.349 | 29 | 176 | 175 | Bacteria | Proteobacteria | Epsilonproteobacteria | Campylobacterales | Campylobacteraceae |
| *Aromatoleum* | 0.541 | 0.278 | 0.473 | 59 | 210 | 78 | Bacteria | Proteobacteria | Betaproteobacteria | Rhodocyclales | Rhodocyclaceae |
| *Asticcacaulis* | 0.352 | 0.910 | 0.437 | 155 | 1 | 103 | Bacteria | Proteobacteria | Alphaproteobacteria | Caulobacterales | Caulobacteraceae |
| *Aurantimonas* | 0.284 | 0.538 | 0.485 | 204 | 69 | 64 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Aurantimonadaceae |
| *Azoarcus* | 0.498 | 0.292 | 0.473 | 78 | 205 | 76 | Bacteria | Proteobacteria | Betaproteobacteria | Rhodocyclales | Rhodocyclaceae |
| *Azorhizobium* | 0.413 | 0.348 | 0.388 | 122 | 166 | 136 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae |
| *Azospirillum* | 0.416 | 0.420 | 0.388 | 120 | 126 | 134 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodospirillales | Rhodospirillaceae |
| *Azotobacter* | 0.450 | 0.473 | 0.588 | 106 | 99 | 31 | Bacteria | Proteobacteria | Gammaproteobacteria | Pseudomonadales | Pseudomonadaceae |
| *Bacteroides* | 0.472 | 0.317 | 0.343 | 91 | 188 | 184 | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Bacteroidaceae |
| *Bartonella* | 0.461 | 0.423 | 0.337 | 100 | 123 | 191 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Bartonellaceae |
| *Beijerinckia* | 0.255 | 0.341 | 0.388 | 224 | 172 | 137 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Blattabacterium* | 0.683 | 0.383 | 0.339 | 17 | 151 | 189 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Blattabacteriaceae |
| *Bordetella* | 0.423 | 0.417 | 0.561 | 117 | 128 | 39 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Alcaligenaceae |
| *Bradyrhizobium* | 0.556 | 0.748 | 0.481 | 52 | 36 | 70 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Bradyrhizobiaceae |
| *Brucella* | 0.358 | 0.414 | 0.330 | 148 | 131 | 199 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Brucellaceae |
| *Burkholderia* | 0.293 | 0.384 | 0.783 | 201 | 149 | 4 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Burkholderiaceae |
| *Candidatus Accumulibacter* | 0.567 | 0.278 | 0.473 | 48 | 209 | 77 | Bacteria | Proteobacteria | Betaproteobacteria | unclassified (derived from Betaproteobacteria) | unclassified (derived from Betaproteobacteria) |
| *Candidatus Amoebophilus* | 0.541 | 0.333 | 0.334 | 58 | 175 | 192 | Bacteria | Bacteroidetes | unclassified (derived from Bacteroidetes) | unclassified (derived from Bacteroidetes) | unclassified (derived from Bacteroidetes) |
| *Candidatus Azobacteroides* | 0.683 | 0.444 | 0.299 | 18 | 108 | 220 | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | unclassified (derived from Bacteroidales) |
| *Candidatus Liberibacter* | 0.399 | 0.274 | 0.514 | 128 | 213 | 57 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiaceae |
| *Candidatus Protochlamydia* | 0.664 | 0.473 | 0.392 | 24 | 98 | 128 | Bacteria | Chlamydiae | Chlamydiae (class) | Chlamydiales | Parachlamydiaceae |
| *Candidatus Sulcia* | 0.528 | 0.420 | 0.384 | 63 | 127 | 144 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | unclassified (derived from Flavobacteriales) |
| *Capnocytophaga* | 0.538 | 0.435 | 0.339 | 62 | 118 | 188 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Catenibacterium* | 0.639 | 0.269 | 0.276 | 31 | 214 | 230 | Bacteria | Firmicutes | Erysipelotrichi | Erysipelotrichales | Erysipelotrichaceae |
| *Catenulispora* | 0.347 | 0.811 | 0.370 | 158 | 16 | 161 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Catenulisporaceae |
| *Cellvibrio* | 0.448 | 0.306 | 0.658 | 109 | 198 | 16 | Bacteria | Proteobacteria | Gammaproteobacteria | Pseudomonadales | Pseudomonadaceae |
| *Chelativorans* | 0.315 | 0.314 | 0.333 | 179 | 190 | 198 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Phyllobacteriaceae |
| *Chitinophaga* | 0.461 | 0.752 | 0.377 | 101 | 29 | 149 | Bacteria | Bacteroidetes | Sphingobacteria | Sphingobacteriales | unclassified (derived from Sphingobacteriales) |
| *Chryseobacterium* | 0.700 | 0.444 | 0.328 | 9 | 107 | 201 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Citreicella* | 0.270 | 0.595 | 0.533 | 217 | 51 | 46 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Citrobacter* | 0.299 | 0.257 | 0.673 | 196 | 221 | 11 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Citromicrobium* | 0.281 | 0.783 | 0.306 | 207 | 24 | 217 | Bacteria | Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae |
| *Colwellia* | 0.476 | 0.313 | 0.389 | 88 | 192 | 130 | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | Colwelliaceae |
| *Comamonas* | 0.522 | 0.590 | 0.452 | 68 | 52 | 93 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Comamonadaceae |
| *Congregibacter* | 0.400 | 0.388 | 0.601 | 127 | 146 | 28 | Bacteria | Proteobacteria | Gammaproteobacteria | unclassified (derived from Gammaproteobacteria) | unclassified (derived from Gammaproteobacteria) |
| *Corynebacterium* | 0.317 | 0.750 | 0.405 | 178 | 33 | 123 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Corynebacteriaceae |
| *Croceibacter* | 0.683 | 0.482 | 0.385 | 19 | 95 | 141 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Cronobacter* | 0.257 | 0.238 | 0.789 | 223 | 226 | 3 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Cupriavidus* | 0.423 | 0.473 | 0.533 | 118 | 100 | 47 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Burkholderiaceae |
| *Cyanobium* | 0.352 | 0.295 | 0.339 | 156 | 203 | 186 | Bacteria | Cyanobacteria | unclassified (derived from Cyanobacteria) | Chroococcales | unclassified (derived from Chroococcales) |
| *Cytophaga* | 0.667 | 0.518 | 0.355 | 23 | 85 | 173 | Bacteria | Bacteroidetes | Cytophagia | Cytophagales | Cytophagaceae |
| *Dechloromonas* | 0.487 | 0.332 | 0.467 | 82 | 179 | 82 | Bacteria | Proteobacteria | Betaproteobacteria | Rhodocyclales | Rhodocyclaceae |
| *Delftia* | 0.492 | 0.556 | 0.464 | 80 | 63 | 83 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Comamonadaceae |
| *Dermacoccus* | 0.411 | 0.848 | 0.347 | 123 | 8 | 181 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Dermacoccaceae |
| *Dickeya* | 0.590 | 0.499 | 0.673 | 42 | 88 | 9 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Dinoroseobacter* | 0.255 | 0.615 | 0.497 | 228 | 46 | 61 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Dokdonia* | 0.528 | 0.343 | 0.431 | 64 | 170 | 105 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Dyadobacter* | 0.590 | 0.369 | 0.229 | 44 | 157 | 233 | Bacteria | Bacteroidetes | Cytophagia | Cytophagales | Cytophagaceae |
| *Edwardsiella* | 0.382 | 0.229 | 0.520 | 136 | 231 | 53 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Eggerthella* | 0.338 | 0.211 | 0.313 | 165 | 233 | 206 | Bacteria | Actinobacteria | Actinobacteria (class) | Coriobacteriales | Coriobacteriaceae |
| *Ehrlichia* | 0.297 | 0.295 | 0.383 | 197 | 202 | 145 | Bacteria | Proteobacteria | Alphaproteobacteria | Rickettsiales | Anaplasmataceae |
| *Enhydrobacter* | 0.696 | 0.530 | 0.425 | 12 | 78 | 112 | Bacteria | Proteobacteria | Gammaproteobacteria | Pseudomonadales | Moraxellaceae |
| *Enterobacter* | 0.483 | 0.449 | 0.561 | 83 | 105 | 38 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Erwinia* | 0.518 | 0.406 | 0.590 | 70 | 137 | 30 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Erythrobacter* | 0.281 | 0.783 | 0.334 | 208 | 25 | 196 | Bacteria | Proteobacteria | Alphaproteobacteria | Sphingomonadales | Erythrobacteraceae |
| *Escherichia* | 0.504 | 0.404 | 0.771 | 73 | 138 | 5 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Filifactor* | 0.234 | 0.325 | 0.205 | 233 | 185 | 234 | Bacteria | Firmicutes | Clostridia | Clostridiales | Peptostreptococcaceae |
| *Flavobacterium* | 0.538 | 0.518 | 0.328 | 61 | 87 | 200 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Fluoribacter* | 0.195 | 0.406 | 0.334 | 234 | 136 | 197 | Bacteria | Proteobacteria | Gammaproteobacteria | Legionellales | Legionellaceae |
| *Frankia* | 0.378 | 0.811 | 0.373 | 140 | 15 | 152 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Frankiaceae |
| *Fulvimarina* | 0.319 | 0.497 | 0.485 | 177 | 89 | 65 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Aurantimonadaceae |
| *Gemella* | 0.357 | 0.240 | 0.280 | 151 | 225 | 228 | Bacteria | Firmicutes | Bacilli | Bacillales | unclassified (derived from Bacillales) |
| *Geodermatophilus* | 0.306 | 0.750 | 0.357 | 189 | 34 | 170 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Geodermatophilaceae |
| *Glaciecola* | 0.689 | 0.430 | 0.357 | 13 | 121 | 172 | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | Alteromonadaceae |
| *Gordonia* | 0.449 | 0.870 | 0.334 | 107 | 2 | 193 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Gordoniaceae |
| *Gramella* | 0.688 | 0.347 | 0.380 | 15 | 167 | 147 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Hahella* | 0.473 | 0.290 | 0.476 | 89 | 206 | 74 | Bacteria | Proteobacteria | Gammaproteobacteria | Oceanospirillales | Hahellaceae |
| *Halothiobacillus* | 0.370 | 0.227 | 0.454 | 143 | 232 | 91 | Bacteria | Proteobacteria | Gammaproteobacteria | Chromatiales | Halothiobacillaceae |
| *Hartmannella* | 0.628 | 0.328 | 0.405 | 33 | 182 | 122 | Eukaryota | unclassified (derived from Eukaryota) | unclassified (derived from Eukaryota) | unclassified (derived from Eukaryota) | Hartmannellidae |
| *Herbaspirillum* | 0.357 | 0.371 | 0.499 | 150 | 155 | 59 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Oxalobacteraceae |
| *Herminiimonas* | 0.278 | 0.367 | 0.538 | 211 | 158 | 42 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Oxalobacteraceae |
| *Hirschia* | 0.313 | 0.655 | 0.413 | 184 | 44 | 116 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Hyphomonadaceae |
| *Hoeflea* | 0.328 | 0.456 | 0.388 | 170 | 103 | 133 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Phyllobacteriaceae |
| *Holdemania* | 0.269 | 0.269 | 0.258 | 219 | 217 | 231 | Bacteria | Firmicutes | Erysipelotrichi | Erysipelotrichales | Erysipelotrichaceae |
| *Hyphomicrobium* | 0.295 | 0.483 | 0.293 | 199 | 94 | 223 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Hyphomicrobiaceae |
| *Hyphomonas* | 0.261 | 0.770 | 0.370 | 220 | 27 | 163 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Hyphomonadaceae |
| *Intrasporangium* | 0.425 | 0.811 | 0.370 | 115 | 12 | 160 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Intrasporangiaceae |
| *Janibacter* | 0.306 | 0.770 | 0.370 | 190 | 26 | 162 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Intrasporangiaceae |
| *Jannaschia* | 0.355 | 0.434 | 0.482 | 152 | 119 | 69 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Janthinobacterium* | 0.381 | 0.372 | 0.651 | 137 | 153 | 19 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Oxalobacteraceae |
| *Ketogulonicigenium* | 0.500 | 0.560 | 0.604 | 75 | 59 | 26 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Klebsiella* | 0.510 | 0.487 | 0.525 | 72 | 92 | 50 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Kordia* | 0.683 | 0.444 | 0.380 | 20 | 109 | 146 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Kribbella* | 0.404 | 0.750 | 0.510 | 125 | 31 | 58 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Nocardioidaceae |
| *Labrenzia* | 0.355 | 0.483 | 0.388 | 153 | 93 | 131 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Leadbetterella* | 0.596 | 0.461 | 0.364 | 40 | 102 | 167 | Bacteria | Bacteroidetes | Cytophagia | Cytophagales | Cytophagaceae |
| *Leeuwenhoekiella* | 0.585 | 0.202 | 0.431 | 46 | 234 | 107 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Leptolyngbya* | 0.237 | 0.245 | 0.296 | 232 | 224 | 221 | Bacteria | Cyanobacteria | unclassified (derived from Cyanobacteria) | Oscillatoriales | unclassified (derived from Oscillatoriales) |
| *Leptospira* | 0.565 | 0.397 | 0.316 | 49 | 141 | 205 | Bacteria | Spirochaetes | Spirochaetes (class) | Spirochaetales | Leptospiraceae |
| *Leptothrix* | 0.482 | 0.536 | 0.441 | 86 | 72 | 101 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | unclassified (derived from Burkholderiales) |
| *Limnobacter* | 0.346 | 0.328 | 0.538 | 160 | 180 | 44 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Burkholderiaceae |
| *Listonella* | 0.259 | 0.799 | 0.434 | 221 | 22 | 104 | Bacteria | Proteobacteria | Gammaproteobacteria | Vibrionales | Vibrionaceae |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Loktanella* | 0.402 | 0.611 | 0.369 | 126 | 47 | 164 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Magnetospirillum* | 0.495 | 0.420 | 0.387 | 79 | 125 | 140 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodospirillales | Rhodospirillaceae |
| *Mannheimia* | 0.343 | 0.231 | 0.520 | 162 | 229 | 52 | Bacteria | Proteobacteria | Gammaproteobacteria | Pasteurellales | Pasteurellaceae |
| *Maribacter* | 0.546 | 0.317 | 0.431 | 57 | 187 | 106 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Maricaulis* | 0.274 | 0.585 | 0.391 | 214 | 54 | 129 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Hyphomonadaceae |
| *Marinobacter* | 0.561 | 0.310 | 0.527 | 50 | 195 | 49 | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | Alteromonadaceae |
| *Marinomonas* | 0.638 | 0.403 | 0.538 | 32 | 139 | 45 | Bacteria | Proteobacteria | Gammaproteobacteria | Oceanospirillales | Oceanospirillaceae |
| *Maritimibacter* | 0.314 | 0.531 | 0.347 | 181 | 74 | 180 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Marivirga* | 0.620 | 0.357 | 0.351 | 34 | 160 | 174 | Bacteria | Bacteroidetes | Cytophagia | Cytophagales | Flammeovirgaceae |
| *Mesorhizobium* | 0.461 | 0.577 | 0.479 | 102 | 56 | 72 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Phyllobacteriaceae |
| *Methylibium* | 0.487 | 0.536 | 0.441 | 81 | 71 | 100 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | unclassified (derived from Burkholderiales) |
| *Methylobacterium* | 0.285 | 0.446 | 0.457 | 203 | 106 | 90 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Methylobacteriaceae |
| *Methylocella* | 0.292 | 0.406 | 0.444 | 202 | 135 | 95 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae |
| *Methylosinus* | 0.284 | 0.328 | 0.388 | 205 | 183 | 138 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Methylocystaceae |
| *Methylotenera* | 0.464 | 0.367 | 0.359 | 97 | 159 | 169 | Bacteria | Proteobacteria | Betaproteobacteria | Methylophilales | Methylophilaceae |
| *Microcystis* | 0.473 | 0.262 | 0.345 | 90 | 219 | 182 | Bacteria | Cyanobacteria | unclassified (derived from Cyanobacteria) | Chroococcales | unclassified (derived from Chroococcales) |
| *Micromonospora* | 0.313 | 0.727 | 0.373 | 185 | 38 | 156 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Micromonosporaceae |
| *Microscilla* | 0.656 | 0.317 | 0.339 | 28 | 186 | 190 | Bacteria | Bacteroidetes | Cytophagia | Cytophagales | Cytophagaceae |
| *Moraxella* | 0.326 | 0.349 | 0.585 | 173 | 165 | 33 | Bacteria | Proteobacteria | Gammaproteobacteria | Pseudomonadales | Moraxellaceae |
| *Moritella* | 0.791 | 0.454 | 0.483 | 3 | 104 | 68 | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | Moritellaceae |
| *Mucilaginibacter* | 0.608 | 0.558 | 0.372 | 35 | 62 | 158 | Bacteria | Bacteroidetes | Sphingobacteria | Sphingobacteriales | Sphingobacteriaceae |
| *Myxococcus* | 0.504 | 0.373 | 0.295 | 74 | 152 | 222 | Bacteria | Proteobacteria | Deltaproteobacteria | Myxococcales | Myxococcaceae |
| *Nakamurella* | 0.345 | 0.811 | 0.306 | 161 | 17 | 214 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Nakamurellaceae |
| *Neisseria* | 0.464 | 0.302 | 0.429 | 98 | 200 | 110 | Bacteria | Proteobacteria | Betaproteobacteria | Neisseriales | Neisseriaceae |
| *Neorickettsia* | 0.272 | 0.338 | 0.375 | 216 | 173 | 150 | Bacteria | Proteobacteria | Alphaproteobacteria | Rickettsiales | Anaplasmataceae |
| *Nitrobacter* | 0.313 | 0.531 | 0.458 | 186 | 75 | 88 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Bradyrhizobiaceae |
| *Nocardia* | 0.384 | 0.811 | 0.306 | 133 | 13 | 213 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Nocardiaceae |
| *Nocardioides* | 0.378 | 0.811 | 0.348 | 139 | 14 | 177 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Nocardioidaceae |
| *Nocardiopsis* | 0.582 | 0.760 | 0.393 | 47 | 28 | 126 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Nocardiopsaceae |
| *Novosphingobium* | 0.281 | 0.808 | 0.306 | 209 | 21 | 215 | Bacteria | Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae |
| *Oceanibulbus* | 0.371 | 0.744 | 0.485 | 142 | 37 | 63 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Oceanicaulis* | 0.246 | 0.722 | 0.364 | 231 | 40 | 166 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Hyphomonadaceae |
| *Oceanicola* | 0.314 | 0.606 | 0.525 | 182 | 49 | 51 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Ochrobactrum* | 0.369 | 0.552 | 0.481 | 145 | 64 | 71 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Brucellaceae |
| *Octadecabacter* | 0.328 | 0.573 | 0.464 | 171 | 58 | 85 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Oligotropha* | 0.334 | 0.487 | 0.485 | 166 | 91 | 66 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Bradyrhizobiaceae |
| *Orientia* | 0.385 | 0.393 | 0.411 | 132 | 143 | 119 | Bacteria | Proteobacteria | Alphaproteobacteria | Rickettsiales | Rickettsiaceae |
| *Paludibacter* | 0.527 | 0.444 | 0.348 | 66 | 112 | 176 | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Porphyromonadaceae |
| *Pantoea* | 0.549 | 0.444 | 0.576 | 55 | 113 | 34 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Parabacteroides* | 0.471 | 0.261 | 0.343 | 93 | 220 | 185 | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Porphyromonadaceae |
| *Parachlamydia* | 0.419 | 0.480 | 0.369 | 119 | 96 | 165 | Bacteria | Chlamydiae | Chlamydiae (class) | Chlamydiales | Parachlamydiaceae |
| *Paracoccus* | 0.358 | 0.672 | 0.518 | 149 | 42 | 54 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Parascardovia* | 0.599 | 0.560 | 0.370 | 38 | 61 | 159 | Bacteria | Actinobacteria | Actinobacteria (class) | Bifidobacteriales | Bifidobacteriaceae |
| *Parvibaculum* | 0.383 | 0.538 | 0.467 | 134 | 68 | 81 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Phyllobacteriaceae |
| *Parvularcula* | 0.270 | 0.497 | 0.410 | 218 | 90 | 120 | Bacteria | Proteobacteria | Alphaproteobacteria | Parvularculales | Parvularculaceae |
| *Pectobacterium* | 0.483 | 0.437 | 0.651 | 84 | 117 | 18 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Pedobacter* | 0.595 | 0.518 | 0.348 | 41 | 86 | 178 | Bacteria | Bacteroidetes | Sphingobacteria | Sphingobacteriales | Sphingobacteriaceae |
| *Pelagibaca* | 0.273 | 0.585 | 0.514 | 215 | 53 | 56 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Phaeobacter* | 0.334 | 0.560 | 0.418 | 167 | 60 | 113 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Photobacterium* | 0.448 | 0.275 | 0.576 | 110 | 211 | 35 | Bacteria | Proteobacteria | Gammaproteobacteria | Vibrionales | Vibrionaceae |
| *Photorhabdus* | 0.700 | 0.354 | 0.665 | 10 | 164 | 14 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Plesiocystis* | 0.347 | 0.264 | 0.305 | 157 | 218 | 218 | Bacteria | Proteobacteria | Deltaproteobacteria | Myxococcales | Nannocystaceae |
| *Polaribacter* | 0.683 | 0.444 | 0.328 | 21 | 110 | 202 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Polaromonas* | 0.661 | 0.620 | 0.441 | 26 | 45 | 97 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Comamonadaceae |
| *Polynucleobacter* | 0.528 | 0.444 | 0.540 | 65 | 114 | 41 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Burkholderiaceae |
| *Prevotella* | 0.641 | 0.306 | 0.385 | 30 | 196 | 142 | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Prevotellaceae |
| *Propionibacterium* | 0.299 | 0.398 | 0.442 | 195 | 140 | 96 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Propionibacteriaceae |
| *Proteus* | 0.369 | 0.274 | 0.599 | 146 | 212 | 29 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Providencia* | 0.315 | 0.294 | 0.661 | 180 | 204 | 15 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Pseudoalteromonas* | 0.452 | 0.333 | 0.538 | 105 | 177 | 43 | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | Pseudoalteromonadaceae |
| *Pseudovibrio* | 0.467 | 0.341 | 0.327 | 95 | 171 | 204 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Psychrobacter* | 0.383 | 0.290 | 0.668 | 135 | 207 | 12 | Bacteria | Proteobacteria | Gammaproteobacteria | Pseudomonadales | Moraxellaceae |
| *Psychroflexus* | 0.683 | 0.576 | 0.287 | 22 | 57 | 226 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Psychromonas* | 0.697 | 0.326 | 0.567 | 11 | 184 | 37 | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | Psychromonadaceae |
| *Ralstonia* | 0.277 | 0.410 | 0.770 | 212 | 133 | 6 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Burkholderiaceae |
| *Reinekea* | 0.730 | 0.354 | 0.452 | 6 | 163 | 92 | Bacteria | Proteobacteria | Gammaproteobacteria | unclassified (derived from Gammaproteobacteria) | unclassified (derived from Gammaproteobacteria) |
| *Rhizobium* | 0.255 | 0.531 | 0.417 | 225 | 77 | 114 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiaceae |
| *Rhodobacter* | 0.393 | 0.531 | 0.412 | 129 | 73 | 118 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Rhodococcus* | 0.340 | 0.811 | 0.334 | 163 | 18 | 195 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Nocardiaceae |
| *Rhodomicrobium* | 0.276 | 0.428 | 0.426 | 213 | 122 | 111 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Hyphomicrobiaceae |
| *Rhodopseudomonas* | 0.355 | 0.541 | 0.458 | 154 | 66 | 87 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Bradyrhizobiaceae |
| *Rhodospirillum* | 0.516 | 0.372 | 0.395 | 71 | 154 | 125 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodospirillales | Rhodospirillaceae |
| *Rickettsia* | 0.334 | 0.328 | 0.311 | 168 | 181 | 209 | Bacteria | Proteobacteria | Alphaproteobacteria | Rickettsiales | Rickettsiaceae |
| *Riemerella* | 0.527 | 0.408 | 0.328 | 67 | 134 | 203 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Roseibium* | 0.281 | 0.388 | 0.306 | 206 | 147 | 210 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Roseobacter* | 0.411 | 0.529 | 0.617 | 124 | 80 | 22 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Roseovarius* | 0.304 | 0.661 | 0.601 | 194 | 43 | 27 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Ruegeria* | 0.362 | 0.522 | 0.469 | 147 | 81 | 80 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Saccharomonospora* | 0.322 | 0.820 | 0.306 | 176 | 10 | 212 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Pseudonocardiaceae |
| *Saccharophagus* | 0.378 | 0.347 | 0.429 | 138 | 169 | 109 | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | Alteromonadaceae |
| *Saccharopolyspora* | 0.306 | 0.750 | 0.373 | 191 | 35 | 155 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Pseudonocardiaceae |
| *Sagittula* | 0.310 | 0.870 | 0.404 | 188 | 4 | 124 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Salmonella* | 0.426 | 0.230 | 0.699 | 114 | 230 | 8 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Segniliparus* | 0.483 | 0.811 | 0.412 | 85 | 11 | 117 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Segniliparaceae |
| *Serratia* | 0.462 | 0.521 | 0.612 | 99 | 82 | 23 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Shewanella* | 0.587 | 0.316 | 0.673 | 45 | 189 | 10 | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | Shewanellaceae |
| *Shigella* | 0.425 | 0.333 | 0.607 | 116 | 178 | 24 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Sinorhizobium* | 0.314 | 0.390 | 0.458 | 183 | 145 | 89 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiaceae |
| *Sphingobacterium* | 0.540 | 0.444 | 0.348 | 60 | 111 | 179 | Bacteria | Bacteroidetes | Sphingobacteria | Sphingobacteriales | Sphingobacteriaceae |
| *Sphingobium* | 0.347 | 0.856 | 0.430 | 159 | 5 | 108 | Bacteria | Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae |
| *Sphingomonas* | 0.293 | 0.856 | 0.372 | 200 | 6 | 157 | Bacteria | Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae |
| *Spirosoma* | 0.558 | 0.281 | 0.252 | 51 | 208 | 232 | Bacteria | Bacteroidetes | Cytophagia | Cytophagales | Cytophagaceae |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Stackebrandtia* | 0.377 | 0.697 | 0.410 | 141 | 41 | 121 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Glycomycetaceae |
| *Starkeya* | 0.328 | 0.269 | 0.388 | 172 | 216 | 139 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae |
| *Stenotrophomonas* | 0.771 | 0.539 | 0.606 | 4 | 67 | 25 | Bacteria | Proteobacteria | Gammaproteobacteria | Xanthomonadales | Xanthomonadaceae |
| *Stigmatella* | 0.416 | 0.422 | 0.290 | 121 | 124 | 225 | Bacteria | Proteobacteria | Deltaproteobacteria | Myxococcales | Cystobacteraceae |
| *Streptomyces* | 0.306 | 0.811 | 0.373 | 192 | 19 | 153 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Streptomycetaceae |
| *Streptosporangium* | 0.466 | 0.750 | 0.373 | 96 | 30 | 154 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Streptosporangiaceae |
| *Subdoligranulum* | 0.607 | 0.313 | 0.379 | 36 | 194 | 148 | Bacteria | Firmicutes | Clostridia | Clostridiales | Ruminococcaceae |
| *Sulfitobacter* | 0.325 | 0.599 | 0.448 | 174 | 50 | 94 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Teredinibacter* | 0.441 | 0.387 | 0.415 | 113 | 148 | 115 | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | unclassified (derived from Alteromonadales) |
| *Thalassobium* | 0.334 | 0.722 | 0.313 | 169 | 39 | 207 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodobacterales | Rhodobacteraceae |
| *Thauera* | 0.499 | 0.334 | 0.473 | 77 | 174 | 75 | Bacteria | Proteobacteria | Betaproteobacteria | Rhodocyclales | Rhodocyclaceae |
| *Thermomonospora* | 0.304 | 0.811 | 0.464 | 193 | 20 | 84 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Thermomonosporaceae |
| *Thiomonas* | 0.500 | 0.444 | 0.485 | 76 | 115 | 67 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | unclassified (derived from Burkholderiales) |
| *Tolumonas* | 0.601 | 0.314 | 0.618 | 37 | 191 | 21 | Bacteria | Proteobacteria | Gammaproteobacteria | Aeromonadales | Aeromonadaceae |
| *Tsukamurella* | 0.449 | 0.870 | 0.334 | 108 | 3 | 194 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Tsukamurellaceae |
| *Variovorax* | 0.661 | 0.536 | 0.459 | 27 | 70 | 86 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Comamonadaceae |
| *Verminephrobacter* | 0.552 | 0.608 | 0.531 | 54 | 48 | 48 | Bacteria | Proteobacteria | Betaproteobacteria | Burkholderiales | Comamonadaceae |
| *Vibrio* | 0.459 | 0.257 | 0.668 | 104 | 222 | 13 | Bacteria | Proteobacteria | Gammaproteobacteria | Vibrionales | Vibrionaceae |
| *Waddlia* | 0.441 | 0.393 | 0.361 | 112 | 144 | 168 | Bacteria | Chlamydiae | Chlamydiae (class) | Chlamydiales | Waddliaceae |
| *Weissella* | 0.443 | 0.269 | 0.313 | 111 | 215 | 208 | Bacteria | Firmicutes | Bacilli | Lactobacillales | Leuconostocaceae |
| *Wolbachia* | 0.747 | 0.356 | 0.393 | 5 | 162 | 127 | Bacteria | Proteobacteria | Alphaproteobacteria | Rickettsiales | Anaplasmataceae |
| *Xanthobacter* | 0.255 | 0.415 | 0.388 | 226 | 130 | 135 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae |
| *Xanthomonas* | 0.910 | 0.431 | 0.585 | 1 | 120 | 32 | Bacteria | Proteobacteria | Gammaproteobacteria | Xanthomonadales | Xanthomonadaceae |
| *Xenorhabdus* | 0.599 | 0.232 | 0.724 | 39 | 228 | 7 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Xylella* | 0.868 | 0.313 | 0.476 | 2 | 193 | 73 | Bacteria | Proteobacteria | Gammaproteobacteria | Xanthomonadales | Xanthomonadaceae |
| *Yersinia* | 0.481 | 0.306 | 1.000 | 87 | 197 | 1 | Bacteria | Proteobacteria | Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae |
| *Zunongwangia* | 0.688 | 0.347 | 0.343 | 16 | 168 | 183 | Bacteria | Bacteroidetes | Flavobacteria | Flavobacteriales | Flavobacteriaceae |
| *Zymomonas* | 0.251 | 0.856 | 0.440 | 229 | 7 | 102 | Bacteria | Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae |

**Table D.2. Classified bacterial and 2 amoeba genera negatively associated with 3 OP containing genera.** Association strength was calculated via Maximal Information Coefficient (MIC) using representative hits. Sign of association (positive or negative) was determined through Pearson Correlation Coefficient. Rank was based on MIC scores across all genera in the table. *Legionella*, *Mycobacterium* and *Pseudomonas* has been abbreviated as "*Leg.*", "*Myco.*" and "*Pseu.*", respectively.

| Genus | MIC scores with | | | Ranked by MIC score | | | Domain | Phylum | Class | Order | Family |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Leg.* | *Myco.* | *Pseu.* | *Leg.* | *Myco.* | *Pseu.* | | | | | |
| Acetohalobium | -0.525 | -0.549 | -0.229 | 21 | 3 | 67 | Bacteria | Firmicutes | Clostridia | Halanaerobiales | Halobacteroidaceae |
| Acidaminococcus | -0.421 | -0.433 | -0.380 | 42 | 26 | 26 | Bacteria | Firmicutes | Negativicutes | Selenomonadales | Acidaminococcaceae |
| Acidobacterium | -0.732 | -0.447 | -0.307 | 1 | 23 | 47 | Bacteria | Acidobacteria | Acidobacteria (class) | Acidobacteriales | Acidobacteriaceae |
| Acidothermus | -0.444 | -0.285 | -0.243 | 36 | 65 | 64 | Bacteria | Actinobacteria | Actinobacteria (class) | Actinomycetales | Acidothermaceae |
| Alicyclobacillus | -0.532 | -0.376 | -0.298 | 18 | 42 | 53 | Bacteria | Firmicutes | Bacilli | Bacillales | Alicyclobacillaceae |
| Ammonifex | -0.602 | -0.461 | -0.235 | 11 | 19 | 66 | Bacteria | Firmicutes | Clostridia | Thermoanaerobacterales | Thermoanaerobacteraceae |
| Anaerotruncus | -0.389 | -0.216 | -0.370 | 56 | 67 | 27 | Bacteria | Firmicutes | Clostridia | Clostridiales | Ruminococcaceae |
| Anoxybacillus | -0.395 | -0.328 | -0.459 | 49 | 59 | 9 | Bacteria | Firmicutes | Bacilli | Bacillales | Bacillaceae |
| Aquifex | -0.413 | -0.481 | -0.404 | 47 | 13 | 18 | Bacteria | Aquificae | Aquificae (class) | Aquificales | Aquificaceae |
| Bulleidia | -0.331 | -0.420 | -0.370 | 64 | 33 | 28 | Bacteria | Firmicutes | Erysipelotrichi | Erysipelotrichales | Erysipelotrichaceae |
| Caldanaerobacter | -0.395 | -0.370 | -0.459 | 50 | 46 | 6 | Bacteria | Firmicutes | Clostridia | Thermoanaerobacterales | Thermoanaerobacteraceae |
| Caldicellulosiruptor | -0.421 | -0.355 | -0.367 | 43 | 51 | 34 | Bacteria | Firmicutes | Clostridia | Thermoanaerobacterales | Thermoanaerobacterales Family III. Incertae Sedis |
| Caminibacter | -0.410 | -0.366 | -0.382 | 48 | 48 | 25 | Bacteria | Proteobacteria | Epsilonproteobacteria | Nautiliales | Nautiliaceae |
| Candidatus Carsonella | -0.230 | -0.293 | -0.269 | 67 | 63 | 62 | Bacteria | Proteobacteria | Gammaproteobacteria | unclassified (derived from Gammaproteobacteria) | unclassified (derived from Gammaproteobacteria) |
| Candidatus Desulforudis | -0.638 | -0.519 | -0.246 | 8 | 4 | 63 | Bacteria | Firmicutes | Clostridia | Clostridiales | Peptococcaceae |
| Candidatus Koribacter | -0.732 | -0.429 | -0.300 | 2 | 27 | 48 | Bacteria | Acidobacteria | unclassified (derived from Acidobacteria) | unclassified (derived from Acidobacteria) | unclassified (derived from Acidobacteria) |
| Candidatus Solibacter | -0.332 | -0.458 | -0.432 | 63 | 21 | 13 | Bacteria | Acidobacteria | Solibacteres | Solibacterales | Solibacteraceae |
| Carboxydothermus | -0.395 | -0.418 | -0.459 | 51 | 37 | 4 | Bacteria | Firmicutes | Clostridia | Thermoanaerobacterales | Thermoanaerobacteraceae |
| Chlorobaculum | -0.525 | -0.375 | -0.319 | 22 | 44 | 44 | Bacteria | Chlorobi | Chlorobia | Chlorobiales | Chlorobiaceae |
| Chloroflexus | -0.732 | -0.501 | -0.298 | 3 | 7 | 50 | Bacteria | Chloroflexi | Chloroflexi (class) | Chloroflexales | Chloroflexaceae |
| Deferribacter | -0.429 | -0.427 | -0.278 | 38 | 28 | 60 | Bacteria | Deferribacteres | Deferribacteres (class) | Deferribacterales | Deferribacteraceae |
| Dehalococcoides | -0.641 | -0.347 | -0.368 | 5 | 53 | 31 | Bacteria | Chloroflexi | Dehalococcoidetes | unclassified (derived from Dehalococcoidetes) | unclassified (derived from Dehalococcoidetes) |
| Dehalogenimonas | -0.507 | -0.347 | -0.280 | 23 | 54 | 58 | Bacteria | Chloroflexi | Dehalococcoidetes | unclassified (derived from Dehalococcoidetes) | unclassified (derived from Dehalococcoidetes) |
| Dethiobacter | -0.373 | -0.375 | -0.329 | 58 | 43 | 41 | Bacteria | Firmicutes | Clostridia | Clostridiales | Syntrophomonadaceae |
| Dialister | -0.371 | -0.334 | -0.345 | 59 | 58 | 38 | Bacteria | Firmicutes | Negativicutes | Selenomonadales | Veillonellaceae |
| Dictyoglomus | -0.579 | -0.425 | -0.313 | 12 | 29 | 45 | Bacteria | Dictyoglomi | Dictyoglomia | Dictyoglomales | Dictyoglomaceae |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ethanoligenens | -0.363 | -0.317 | -0.319 | 60 | 61 | 43 | Bacteria | Firmicutes | Clostridia | Clostridiales | Ruminococcaceae |
| Geobacillus | -0.531 | -0.483 | -0.411 | 20 | 12 | 15 | Bacteria | Firmicutes | Bacilli | Bacillales | Bacillaceae |
| Halothermothrix | -0.421 | -0.373 | -0.404 | 44 | 45 | 22 | Bacteria | Firmicutes | Clostridia | Halanaerobiales | Halanaerobiaceae |
| Heliobacterium | -0.363 | -0.369 | -0.367 | 61 | 47 | 33 | Bacteria | Firmicutes | Clostridia | Clostridiales | Heliobacteriaceae |
| Herpetosiphon | -0.313 | -0.420 | -0.459 | 66 | 34 | 3 | Bacteria | Chloroflexi | Chloroflexi (class) | Herpetosiphonales | Herpetosiphonaceae |
| Hydrogenobacter | -0.460 | -0.510 | -0.441 | 34 | 6 | 12 | Bacteria | Aquificae | Aquificae (class) | Aquificales | Aquificaceae |
| Ktedonobacter | -0.579 | -0.328 | -0.459 | 13 | 60 | 10 | Bacteria | Chloroflexi | Ktedonobacteria | Ktedonobacterales | Ktedonobacteraceae |
| Marinococcus | -0.183 | -0.183 | -0.183 | 68 | 68 | 68 | Bacteria | Firmicutes | Bacilli | Bacillales | Bacillaceae |
| Moorella | -0.579 | -0.425 | -0.298 | 14 | 30 | 52 | Bacteria | Firmicutes | Clostridia | Thermoanaerobacterales | Thermoanaerobacteraceae |
| Nautilia | -0.426 | -0.499 | -0.359 | 39 | 8 | 36 | Bacteria | Proteobacteria | Epsilonproteobacteria | Nautiliales | Nautiliaceae |
| Nostoc | -0.421 | -0.358 | -0.426 | 45 | 50 | 14 | Bacteria | Cyanobacteria | unclassified (derived from Cyanobacteria) | Nostocales | Nostocaceae |
| Oceanithermus | -0.625 | -0.711 | -0.313 | 9 | 1 | 46 | Bacteria | Deinococcus-Thermus | Deinococci | Thermales | Thermaceae |
| Oscillochloris | -0.483 | -0.483 | -0.279 | 28 | 10 | 59 | Bacteria | Chloroflexi | Chloroflexi (class) | Chloroflexales | Oscillochloridaceae |
| Pediococcus | -0.579 | -0.314 | -0.459 | 15 | 62 | 11 | Bacteria | Firmicutes | Bacilli | Lactobacillales | Lactobacillaceae |
| Pelotomaculum | -0.564 | -0.567 | -0.404 | 16 | 2 | 17 | Bacteria | Firmicutes | Clostridia | Clostridiales | Peptococcaceae |
| Petrotoga | -0.532 | -0.347 | -0.339 | 19 | 52 | 39 | Bacteria | Thermotogae | Thermotogae (class) | Thermotogales | Thermotogaceae |
| Rhodothermus | -0.470 | -0.483 | -0.287 | 32 | 11 | 56 | Bacteria | Bacteroidetes | Sphingobacteria | Sphingobacteriales | Rhodothermaceae |
| Roseiflexus | -0.495 | -0.469 | -0.326 | 26 | 16 | 42 | Bacteria | Chloroflexi | Chloroflexi (class) | Chloroflexales | Chloroflexaceae |
| Roseomonas | -0.498 | -0.225 | -0.337 | 25 | 66 | 40 | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodospirillales | Acetobacteraceae |
| Sphaerobacter | -0.469 | -0.418 | -0.385 | 33 | 36 | 24 | Bacteria | Chloroflexi | Thermomicrobia (class) | Sphaerobacterales | Sphaerobacteraceae |
| Streptobacillus | -0.554 | -0.291 | -0.272 | 17 | 64 | 61 | Bacteria | Fusobacteria | Fusobacteria (class) | Fusobacteriales | Fusobacteriaceae |
| Sulfurihydrogenibium | -0.429 | -0.385 | -0.281 | 37 | 41 | 57 | Bacteria | Aquificae | Aquificae (class) | Aquificales | Hydrogenothermaceae |
| Symbiobacterium | -0.395 | -0.461 | -0.459 | 52 | 20 | 2 | Bacteria | Firmicutes | Clostridia | Clostridiales | Clostridiales Family XVIII. Incertae Sedis |
| Synechococcus | -0.505 | -0.464 | -0.369 | 24 | 18 | 29 | Bacteria | Cyanobacteria | unclassified (derived from Cyanobacteria) | Chroococcales | unclassified (derived from Chroococcales) |
| Syntrophothermus | -0.445 | -0.340 | -0.459 | 35 | 56 | 7 | Bacteria | Firmicutes | Clostridia | Clostridiales | Syntrophomonadaceae |
| Terriglobus | -0.641 | -0.422 | -0.368 | 6 | 31 | 30 | Bacteria | Acidobacteria | Acidobacteria (class) | Acidobacteriales | Acidobacteriaceae |
| Thermaerobacter | -0.395 | -0.480 | -0.459 | 53 | 14 | 1 | Bacteria | Firmicutes | Clostridia | Clostridiales | Clostridiales Family XVII. Incertae Sedis |
| Thermanaerovibrio | -0.474 | -0.364 | -0.293 | 30 | 49 | 55 | Bacteria | Synergistetes | Synergistia | Synergistales | Synergistaceae |
| Thermincola | -0.425 | -0.433 | -0.367 | 40 | 25 | 32 | Bacteria | Firmicutes | Clostridia | Clostridiales | Peptococcaceae |
| Thermoanaerobacter | -0.395 | -0.337 | -0.459 | 54 | 57 | 8 | Bacteria | Firmicutes | Clostridia | Thermoanaerobacterales | Thermoanaerobacteraceae |
| Thermoanaerobacterium | -0.421 | -0.420 | -0.404 | 46 | 32 | 20 | Bacteria | Firmicutes | Clostridia | Thermoanaerobacterales | Thermoanaerobacterales Family III. Incertae Sedis |
| Thermobaculum | -0.620 | -0.473 | -0.298 | 10 | 15 | 51 | Bacteria | unclassified (derived from Bacteria) | unclassified (derived from Bacteria) | unclassified (derived from Bacteria) | unclassified (derived from Bacteria) |
| Thermocrinis | -0.421 | -0.419 | -0.358 | 41 | 35 | 37 | Bacteria | Aquificae | Aquificae (class) | Aquificales | Aquificaceae |
| Thermomicrobium | -0.639 | -0.449 | -0.393 | 7 | 22 | 23 | Bacteria | Chloroflexi | Thermomicrobia (class) | Thermomicrobiales | Thermomicrobiaceae |
| Thermosediminibacter | -0.483 | -0.347 | -0.298 | 29 | 55 | 54 | Bacteria | Firmicutes | Clostridia | Thermoanaerobacterales | Thermoanaerobacterales Family III. Incertae Sedis |
| Thermosinus | -0.348 | -0.417 | -0.241 | 62 | 38 | 65 | Bacteria | Firmicutes | Negativicutes | Selenomonadales | Veillonellaceae |
| Thermosynechococcus | -0.381 | -0.469 | -0.404 | 57 | 17 | 19 | Bacteria | Cyanobacteria | unclassified (derived | Chroococcales | unclassified (derived |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | from Cyanobacteria) | | from Chroococcales) |
| Thermotoga | -0.490 | -0.397 | -0.459 | 27 | 40 | 5 | Bacteria | Thermotogae | Thermotogae (class) | Thermotogales | Thermotogaceae |
| Thermus | -0.732 | -0.514 | -0.298 | 4 | 5 | 49 | Bacteria | Deinococcus-Thermus | Deinococci | Thermales | Thermaceae |
| Trichodesmium | -0.395 | -0.415 | -0.404 | 55 | 39 | 21 | Bacteria | Cyanobacteria | unclassified (derived from Cyanobacteria) | Oscillatoriales | unclassified (derived from Oscillatoriales) |
| Truepera | -0.471 | -0.491 | -0.364 | 31 | 9 | 35 | Bacteria | Deinococcus-Thermus | Deinococci | Deinococcales | Trueperaceae |
| Ureaplasma | -0.313 | -0.440 | -0.409 | 65 | 24 | 16 | Bacteria | Tenericutes | Mollicutes | Mycoplasmatales | Mycoplasmataceae |

# CHAPTER 6. CONCLUSION AND FUTURE WORK

This dissertation contributed to the current understanding on building plumbing microbiome and initiated field attempts towards a 'probiotic' approach for engineering building plumbing microbiome. The results unveiled only the tip of the iceburg, yet the information is clear that the building plumbing microbiome displays significant spatial and temperal heterogeneity. At the geographical scale, source water chemistry and water treatment process appear to be the predominant factor on microbiome composition at the point-of-use. Within a single household, selection on plumbing design (e.g., pipe material) and configuration (e.g., pipe orientation), operation condition (e.g., water heater temperature set point), and user behavior (flush or not flush before usage; water use frequency) can substantially change the microbiomes received at consumer's end. An important theme that these studies collectively support is that a preventive 'probiotic' framework is a promising means of OPs control in the future, as was illustrated in studies of the water heater temperature set point and mild heat shock treatment.

This dissertation also offered fundamental insight into the relationship between the microbial composition of the bulk water and biofilm. It is important to note that there are no ideal methods of sampling biofilms, with each subject to limitations. The specific biofilm sampling protocol employed in this dissertation involved repeated swabbing of the same area after certain time period. This method effectively eliminated the spatial heterogeneity in biofilms across different areas within the pipes, and allowed exploration of the biofilm established under the new condition. From a microbial ecology point of view, it would be of future interest to sampling undisturbed biofilms each time (e.g., cut a section of pipe) to explore the response of established microbiome to the disturbance (i.e., new condition). This is of practical implications as it more closely resembles the response of microbiome to each individual engineering condition.

Another general question of future interest is how to effectively capture the key associations among microbes. Preliminary attempts have been made, including identification of the microbes closely associated with OP containing genera, and association among top abundant bacterial genera. These attempts are on the entire collection of samples tested, thus mostly captures the 'overarching' associations, i.e., strong associations with consistent signs (positive or negative)

under various conditions. In theory, association among microbes would vary under different engineering conditions, which requires further scrutinize within the allowable range of current statistical power. In the long term, data mining remains a technical challenge, and optimization of data visualization is always worth improving.