Profiling of Microbial Communities, Antibiotic Resistance, Functional Genes, and Biodegradable Dissolved Organic Carbon in a Carbon-Based Potable Water Reuse System

Matthew Forrest Blair

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, Chair
Charles B. Bott
Marc A. Edwards
Peter J. Vikesland

January 30th, 2023
Blacksburg VA

Keywords: potable reuse, water reuse, next generation sequencing, antibiotic resistance, functional metagenomic analysis, microbial ecology, advanced water treatment, biodegradable dissolved organic carbon

Profiling of Microbial Communities, Antibiotic Resistance, Functional Genes, and Biodegradable Dissolved Organic Carbon in a Carbon-Based Potable Water Reuse System

Matthew Forrest Blair

ABSTRACT

Water reuse has become a promising alternative to alleviate stress on conventional freshwater resources in the face of population growth, sea level rise, source water depletion, eutrophication of water bodies, and climate change. Potable water reuse intentionally looks to purify wastewater effluent to drinking water quality or better through the development and implementation of advanced treatment trains. While membrane-based treatment has become a widely-adopted treatment step to meet this purpose, there is growing interest in implementing treatment trains that harness microorganisms as a more sustainable and less energy-intensive means of removing contaminants of emerging concern (CECs), through biological degradation or transformation. In this dissertation, various aspects of the operation of a microbially-active carbon-based advanced treatment train producing water intended for potable reuse are examined, including fate of dissolved organic carbon, underlying microbial populations, and functional genes are explored. Further, dynamics associated with antibiotic resistance genes (ARGs), identified as a microbially-relevant CECs, are also assessed. Overall, this dissertation advances understanding associated with the interplay between and within treatment processes as they relate to removal of various organic carbon fractions, microbially community dynamics, functional genes, and ARGs. Further, when relevant, these insights are contextualized to operational conditions, process upsets, water quality parameters, and other intended water uses within the water industry with the goal of broadening the application of advanced molecular tools beyond the scope of academic research.

Specifically, this dissertation illuminates relationships among organic carbon fractions and molecular markers within an advanced treatment train employing flocculation, coagulation, and sedimentation (FlocSed), ozonation, biologically active carbon (BAC) filtration, granular active carbon (GAC) contacting, and UV disinfection. Biodegradable dissolved organic carbon (BDOC) analysis was adapted specifically as an assay relevant to assessing dissolved organic carbon biodegradability by BAC/GAC-biofilms and applied to profile biodegradable/non-biodegradable organic carbon as wastewater effluent passed through each of these treatment stages. Of particular interest was the role of ozonation in producing bioavailable organic carbon that can be effectively removed by BAC filtration. In addition to understanding the removal of fractionalized organic carbon, next generation DNA sequencing technologies (NGS) were utilized to better understand the microbial dynamics characteristic of complex microbial communities during disinfection and biological treatment. Specifically, this analysis was focused on succession and colonization of taxa, genes related to a wide range of functional interests (e.g. metabolic processes, horizontal gene transfer, DNA repair, and nitrogen cycling), and microbial CECs. Finally, NGS technologies were employed to assess the differences between a wide range of water use categories, including conventional drinking water, potable reuse, and non-potable reuse effluent's microbiomes to identify core and discriminatory taxa associated with intended water usage. The outcomes of this dissertation provide valuable information for optimizing carbon-based treatment trains as an alternative to membrane-based treatment for sustainable water reuse and also advance the application of NGS as a diagnostic tool for assessing the efficacy of various water treatment technologies for achieving treatment goals.

Profiling of Microbial Communities, Antibiotic Resistance, Functional Genes, and Biodegradable Dissolved Organic Carbon in a Carbon-Based Potable Water Reuse System

Matthew Forrest Blair

GENERAL AUDIENCE ABSTRACT

Several factors have led to increased stress on conventional drinking water sources and widespread global water scarcity. Projections indicate that continued population growth, increased water demand, and degradation of current freshwater resources will negatively contribute to water needs and underscore the need to secure new potable (i.e. fit for human consumption) sources. Water reuse is a promising alternative to offset the growing demands on traditional potable sources and ameliorate negative consequences associated with water scarcity. Discharge of treated wastewater to marine environments is especially a lost opportunity, as the water will no longer be of value to freshwater habitats or as a drinking water source. Water reuse challenges the conventional wastewater treatment paradigm by providing advanced treatment of wastewater effluent to produce a valuable resource that can be safely used directly for either non-potable (e.g., irrigation, firefighting) or potable (i.e., drinking water) applications.

The means of achieving advanced treatment of wastewater effluents can take many forms, commonly relying on the utilization of membrane filtration. However, membrane filtration is an intensive process and suffers from high initial costs, high operational costs, membrane fouling with time, and the production of a salty and difficult to dispose of waste stream. These drawbacks have motivated the water reuse industry to explore more sustainable approaches to achieving high quality effluents. One such alternative relies on the utilization of microorganisms to provide biological degradation and transformation of contaminants through a process known as biologically active filtration (BAF). Comparatively to membrane systems, BAF is more cost effective and produces significantly fewer byproducts while still producing high quality treated water for reuse. However, the range in quality of the resulting treated water has not yet been fully established, in part due to the lack of understanding of the complex microbial communities responsible for biological treatment.

As water and wastewater treatment technologies have evolved over the past century, many biological treatments have remained largely ‘black box’ due to the lack of effective tools to identify the tens of thousands of species of microbes that inhabit a typical system and to track their dynamics with time. Instead, analysis has largely focused on basic water quality indicators. This dissertation takes important steps in advancing the implementation of the study of DNA and biodegradable organic carbon (BDOC) analysis to improve understanding of the mechanisms that drive different water reuse treatment technologies and to identify potential vulnerabilities. Insights gained through application of these tools are contextualized to observed operational conditions, process upsets, and water quality measurements. This helped to advance the use of DNA-based tools to better inform water treatment engineering practice. Specifically, this dissertation dives into the relationships between organic carbon and DNA-based markers within an advanced treatment train employing flocculation, coagulation, and sedimentation (FlocSed), ozonation, biologically active carbon (BAC) filtration, granular active carbon (GAC) contacting, and UV disinfection.

Development and application of the BDOC test revealed that the bulk of organic carbon entering the treatment train is dissolved. Further, BDOC analysis served to characterize the

impact of specific treatment processes and changes in operational conditions on both biodegradable and non-biodegradable organic carbon fractions. Such information can help to inform continued process optimization.

Utilization of DNA-based technologies shed light on the functional capacity of microbial communities present within each stage of treatment and the fate of antibiotic resistance genes (ARGs). ARGs are of concern because, when present in human pathogens, they can result in the failure of antibiotics to cure deadly infections. Other functional genes of interest were also examined using the DNA-based analysis, including genes driving metabolic processes and nitrogen cycling that are critical to water purification during BAF treatment. Also, the DNA-based analyses made it possible to better understand the effects of disinfectants on microbes. Interestingly, some ARG types increased in relative abundance (a measure analogous to percent composition) response to treatments, such as disinfection, and others decreased.

Characterization of the microbial communities and their dynamic response to changing operation conditions were also observed. For example, it was possible to characterize how the profiles of microbes changed with time, an ecological process called succession, during BAC filtration and GAC contacting. Generally, this analysis, coupled with the functional analysis, shed light on the important, divergent roles of bacterial communities on organic degradation during both BAC and GAC treatment.

Finally, a study was conducted that compared the microbiome (i.e. entire microbial community) between a wide range of conventional drinking water, potable reuse water, and non-potable reuse waters. Here it was found that significant differences existed between the microbial communities of water intended for potable or non-potable usage. This work also looked to expand the application of NGS technologies beyond strictly academic research by developing the application of more advanced DNA-based tools for treatment train assessment and monitoring.

ACKNOWLEDGEMENTS

I want to begin by expressing the sincerest thanks to my primary advisor Dr. Amy Pruden, for her mentorship, assistance, and guidance throughout my time at Virginia Tech. I would like to extend similar well wishes to my advisors Dr. Charles Bott, Dr. Peter Vikesland and Dr. Marc A. Edwards with a special thanks to Dr. Bott for including me on such an awesome research team. All of you have invested time into my personal and professional development, welcomed me into your lab groups, and made me feel at home during graduate school. Your continued dedication to excellence, both scientific and otherwise, is exceptional and motivates me to do better myself. I can safely say I would not be the student, or person I am today without your explicit involvement in my life. I consider myself extremely blessed to know each and everyone of you, and to have been able to work alongside you on projects that I believe better people's lives.

I would also like to acknowledge the financial support I have received throughout my time in graduate school that has made this dissertation possible. Specifically, the contributions provided by the Hampton Roads Sanitation District, the Water Research Foundation, the Bureau of Reclamation, the Virginia Tech Institute for Critical Technology and Applied Science Center for Science and Engineering of the Exposome, the Virginia Tech Graduate School, and the Virginia Tech Department of Civil and Environmental Engineering. Thank you as well to the Charles E. Via family for supporting my research.

To the members of the Pruden, Edwards, and Vikesland groups, both past and present, thank you all for your day-to-day mentorship, and friendship. I have learned so much from each of you, and cherish the experiences we share, and the ability to be a part of such a wonderful group of people.

To the members of the HRSD SWIFT team, thank you so much for making me feel like a part of the family and including me in such an awesome, life changing project. You all embody the reasons I wanted to become an engineer in the first place and gave my work the meaning that I desperately need. You dedication to figuring it out is inspiring, and I will always treasure our friendships.

I would also like to thank all of my teachers, coaches, mentors, and mentees over my life that are too numerous to name. From grade school through graduate school, from sports to band, and everything in between you have all had an active role in making me the person I am today.

Finally, I would like to thank my friends, family, and loved ones with a special thanks to my Mom. Words cannot express how much I appreciate all the support over the years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

OVERVIEW AND RESEARCH MOTIVATION

The Earth is covered in water, to the approximant tune of 71% of the surface and roughly 332.5 million mi^3 [1]. However, most of the Earth's water, roughly 96%, is saline and unsuitable for human consumption and most other applications. Desalination is an extremely costly process[2,3]. Overall, it is estimated that long-term sources of freshwater make up around 2.5% of the total water stock with roughly 68% of that locked up in permeant ice, glaciers, and permafrost[1]. This leaves an estimated 30% of all freshwaters as groundwater and only 0.26% stored in surface waters[1]. Of these primary freshwater sources, most of the global water usage comes from surface waters[4,5], with utilization of groundwater reservoirs limited by accessibility and a host of other challenges[6–8] in most parts of the world. In certain areas around the globe, and within the United States, continued utilization of both surface and groundwater resources coupled with mismanagement of environmental discharges have contributed to source depletion, degradation of water quality, eutrophication of vital water bodies, land subsidence, sea water intrusion, and a host of other issues that are exacerbated by the impact of global climate change. As it stands, scarcity of clean water is a pressing issue for between 47%[9] to 52%[10] of the Earth's population in at-least one month per year with estimates for 2050 approaching 57%[9].

The concerns related to water scarcity, increasing global demand, and limited ability to increase conventional water sources has motivated development of new freshwater sources. Water reuse represents a very promising alternative and is rooted in a changing paradigm associated with conventional wastewater treatment. Thus far, wastewater treatment has largely been focused on providing attenuation of contaminants to environmentally safe standards and largely fails to produce a valuable resource. Within the U.S. alone, it is estimated that 32 billion gallons of wastewater is treated daily[11], of which, only 7-8% is intentionally reused[12] making wastewater a massively untapped resource. However, as discharge standards become stricter, and resources become more limited, there is a strong motivation to harness or recapture the many valuable components that comprise wastewater[13,14]. These include nutrients[15,16], energy[17], bioenergy[18] biosolids/fertilizers[19,20], biogases[21,22], and most of all the water itself[23–25].

This change in approach has led to the acknowledgement of the 'One Water' paradigm, championed by the Water Research Foundation, which roots itself in understanding the interconnectivity of all water related sectors (e.g. water supply, wastewater, reuse, watershed management, stormwater, and energy and resource recovery) with the goal of implementing strategies for long-term, sustainable utilization of finite water resources[26]. Water reuse factors heavily into the 'One Water' paradigm as it bridges the illusory divide between different water sectors. Often unacknowledged, *de-facto* reuse, or the downstream repurposing of wastewater effluents, has occurred throughout history and is becoming more widespread as conventional water sources become increasingly stressed[27]. The classic example of *de facto* reuse is a drinking water treatment plant intake situated downstream from a wastewater treatment plant effluent discharge point. In such a scenario, water reuse is unintentional and thus does not benefit from the potential to plan and implement technologies that ensure the overall safety of the practice. When planed in an intentional matter, water reuse practice typically implements various forms of advanced treatment of wastewater effluents. To be truly sustainable, however, it is best to avoid

investment of more energy and resources necessary to achieve the quality of water needed for the intended application. With this in mind, a "fit-for-use" model has been adopted where the intended application of the reuse water dictates the level of necessary treatment. Reuse applications can be broadly categorized as either non-potable and potable. Potable reuse is typically further classified as either direct, in which advanced treated water is routed directly to the drinking water treatment plant, or indirect, wherein the treated water is first stored in a reservoir or aquifer before being subject to drinking water treatment. Regardless of its form, the specific applications of water reuse have the potential to alleviate demand on more conventional water sources[24,25], decrease nutrient rich discharges to impaired water bodies[28], protect against seawater intrusion[29], and provide other benefits that a conventional wastewater discharge does not[29–35]. The expansion of water reuse applications relies on the continued development of cost-effective tertiary treatment trains that can reliably produce high quality effluents devoid of the hazards associated with wastewater effluents, specifically: nutrients, suspended solids, organic pollutants, pathogens, and contaminants of emerging concern (CECs)[36].

CECs are of particular interest because their human and environmental health impacts are not well defined and thus they are generally not subject to regulation. Many CECs exist as trace organic compounds (TrOCs), which are difficult to remove by conventional treatment processes implemented in drinking water[37–41] and wastewater[42–45] treatment. TrOCs are often difficult to biodegrade because of their complex chemical structure and low concentration, which offer little in terms of a carbon and energy source that can sustain microbial growth. CECs are made up of a wide range of contaminants, ranging from chemical to biological, and are often subdivided into categories including endocrine-disrupting compounds (EDCs), pharmaceuticals (PhACs), personal care products (PCPs), and microbial containments (e.g., viruses, antibiotic resistance genes (ARGs)). Advanced treatment technologies, such as activated carbon, ozonation, ultraviolet irradiation, sonodegradation, and membrane filtration[38,39,46–48], can be effectively applied to remove most CECs. However, removal rates are compound specific and complicated by treatment configuration and operational complexities. The efficacy of wastewater treatment for CEC removal is especially difficult to assess due to the amalgamation of treatment processes, influent variability, and localized operational conditions. The physiochemical properties of the individual CECs (e.g., pKA, charge, size, dimensions, and hydrophobicity) and the highly variable configurations of wastewater treatment (e.g., sludge adsorption, aerobic biodegradation, anaerobic biodegradation, anoxic biodegradation, oxidation by O3/chlorine, solid-liquid separation, disinfection)[49] are primary contributors to the complexity of disentangling these removal mechanisms[50]. The nature of water reuse, where the human-altered water cycle is short-circuited to various degrees, exacerbates the consequences of treatment deficiencies. Eliminating storage in the environment, where natural attenuation can serve as a contaminant barrier, further necessitates the importance of a multi-barrier treatment approach to ensure that public health is adequately protected. Therefore, within water reuse applications there is a need to accurately characterize and understand CEC and organic pollutant removal mechanisms, which will aid in informing more targeted and optimized treatment processes[50]. This will, in turn, support a more tailored approach to achieving the level needed to protect public health, depending on the reuse application.

In most water reuse applications, the default means of achieving high quality effluents rely on the utilization of one or multiple forms of membrane filtration, including: microfiltration, ultrafiltration, nanofiltration, and/or reverse osmosis (RO). RO has largely been considered the

gold standard for advanced treatments, particularly in California where[51–54] RO has found itself integrated into many applications of water reuse around the world[24]. However, RO systems, and other membrane-based systems, suffer from operational challenges[55–59] including high initial and operation costs, and the production of a difficult to dispose of brine waste steam. Membrane-based systems also provide limited contaminant degradation or transformation into less toxic substrates, instead providing removal through separation. Further, RO-based systems can remove too many constituents, making the resulting treated water unfit for aquifer recharge without additional costly adjustment of the water chemistry[60]. These drawbacks leave room for more sustainable approaches with many involving the utilization of microorganisms to provides biological degradation and transformation of CECs and organic pollutants. To provide sufficient protection for human health concerns, these biologically-based systems typically rely on multiple barriers for contaminant removal and disinfection.

One such configuration relies on biologically active filtration (BAF) to provide simultaneous organic degradation and filtration. However, due to the nature of upstream biological removal during wastewater treatment, most of the remaining organic contamination is not amenable to further degradation, at-least under typical operating conditions. Therefore, advanced systems looking to employ more sustainable biological treatment typically couple BAF with an upstream oxidative process, such as ozonation, to increase the bioavailability of remaining organic compounds while simultaneously providing pathogen removal. One such combination of treatment processes is commonly referred to as a "carbon-based" treatment train, as it relies on activated carbon based biologically activated carbon filtration (BAC) and granular activated carbon (GAC) contacting following ozonation. BAC filtration utilizes GAC support media as a means to provide synergistic advantages of high surface area for adsorption and microbial colonization. This interplay between contaminants and microorganisms can enhance degradation of organics and uptake or transformation of nutrients while simultaneously allowing for the development of more complex microbial communities and associated functional pathways. Overall, implementation of ozonation with BAC filtration and GAC contacting has been reported to achieve effluent TrOCs compositions that are comparable to RO and ultrafiltration systems at much lower costs[61].

Thus, carbon-based treatment trains have been identified as a promising alternative to RO and membrane treatment. However, a fundamental understanding of achievable water quality is still needed. In particular, the role of the microbial community in metabolizing organic matter and factors governing the microbiological and organic content of final effluents is needed to advance its full-scale application. However, assessing the role of microbial communities and their key functions associated with drinking water, wastewater, or water reuse is methodologically challenging. Before the development of more comprehensive DNA-based analytical approaches, biological analysis has relied on culture-based methods. These methods have conventionally indexed heavily into pathogen monitoring with little regard for non-pathogenic, complex communities. Culture, though valuable for identifying living organisms, is extremely limited in its ability to characterize complex communities due to the overall diversity and general lack of culturability for most organisms[62–64]. Recent technological advances in molecular techniques (i.e., targeting DNA or RNA) are beginning to provide alternatives to culture-based methods for the water industry. Specific gene target-based analysis, such as quantitative polymerase chain reaction (qPCR), is becoming widely applied and now has standardized methods for fecal indicators that have been adopted by the U.S. EPA[65,66]. Most

recently, next generation DNA sequencing (NGS) technologies (e.g. whole genome sequencing, metagenomic sequencing, metatranscriptomic sequencing, and targeted sequencing of amplified gene regions) have completely transformed environmental microbiological through their ability to indiscriminately characterize the breadth and depth of mixed microbial communities and their functional genes within environmental samples[67]. NGS technologies hold significant promise for informing the optimization of biological-based treatments and serving the needs of an ever changing water industry by providing substantial insights into interplay between treatment processes, operational conditions, water quality parameters and the diverse, expansive microbial populations that contribute to water treatment. Decreasing sequencing costs and development of publicly-available analysis pipelines and annotation databases will further make such tools more accessible. Though not without limitations, NGS technologies hold great potential for illuminating the 'black box' of biologically-based treatments within the water industry, especially as more advanced technologies gain a foothold for water reuse applications.

RESEARCH OBJECTIVES

This research takes advantage of a unique pilot treatment train currently in operation by the Hampton Roads Sanitation District's (HRSD) Sustainable Water Initiative for Tomorrow (SWIFT). The treatment train consists of coagulation-flocculation-sedimentation, ozonation, biologically active carbon filtration, granular activated carbon filtration, and ultraviolet (UV) disinfection treatments to purify wastewater treatment plant (WWTP) effluent to drinking water standards or better. HRSD has taken steps to implement the SWIFT process both at pilot- (4.3 gpm) and demonstration-scale (1MGD), which has enabled a number of ongoing and concurrent research projects aimed at understanding and improving the treatment process under realistic conditions. The research described in this dissertation was coordinated with many of these ongoing projects to support a synergistic understanding of the SWIFT process.

The **overarching objective** of the research described in this dissertation was to characterize the microbial dynamics, both in terms of taxonomic and functional gene composition, occurring in time and space throughout a pilot-scale advanced water treatment train employing ozone-BAC-GAC treatment for potable water reuse. A major theme of the research is advancement of NGS as a practical tool for gaining valuable mechanistic insight into each unit process in order to inform overall improvement of design and operation. Molecular-based insights were benchmarked to operational conditions and water chemistry as well as contextualized to other water types, when possible. Further, the fates of biodegradable and non-biodegradable dissolved organic carbon were characterized throughout each stage of treatment. Together, the research here will help to optimize efficacy of a carbon-based reuse train by characterizing the underlying microbially community compositions and functional profiles responsible for effective treatment performance and mitigating the selection and dissemination of biological CECs (such as ARGs). The specific objectives pursed were to:

1) Develop a biodegradable dissolved organic carbon (BDOC) method that is suitable for biofilm-based water reuse treatment application and apply it towards gaining insight into the efficacy of various unit processes for removing biodegradable versus non-biodegradable organic carbon fractions.
2) Profile ARGs (i.e., the resistome) through various stages of carbon-based advanced water treatment and assess how both intentional changes to process operation and

unintentional upset events influence observed shifts in corresponding resistome characteristics.

3) Utilize NGS to conduct a holistic analysis of the functional capacity of microbial communities to degrade TrOCs and respond to disinfection technologies through various treatment stages of a carbon-based potable reuse train.
4) Characterize microbial colonization and succession of a carbon-based potable reuse train, both spatially and temporally, and relate shifts in microbial community structure to fluctuations in performance and changes in operational conditions and water chemistry.
5) Contextualize water reuse effluent microbial communities relative to those characteristic of other intended water uses as a step towards applying NGS sequencing as a diagnostic tool for categorizing water quality for intended application.

ANNOTATED DISSERTATION OUTLINE AND ATTRIBUTIONS

Chapter 1: *Introduction*

This chapter details the motivation for this research and defines and contextualizes specific research objectives.

Chapter 2: *Biodegradable Dissolved Organic Carbon Profiling Reveals Capacity of Carbon-Based Potable Reuse Treatment*

This chapter consists of a manuscript intended for publication and relates to Research Objective (1). Specifically, Chapter 2 focuses on the development of a BDOC assay appropriate for evaluating biofilm-based water reuse treatment and applying it towards assessing the fate of biodegradable and non-biodegradable dissolved organic carbon throughout a carbon-based advanced treatment train. This research highlights each treatment process's ability to remove various forms of organic carbon over changing operational conditions and water quality parameters while identifying room for process optimization as it relates to removal of organic carbon.

This manuscript has been submitted to *AWWA Water Science* and is currently under revision following a decision of rejection and resubmission.

Attributions: I led protocol refinement, sample collection, in lab analysis, data analysis, and manuscript writing. Ramola Vaidya was primarily responsible for pilot operation under Germano Salazar Benites's direct management. Larry Schimmoller and Tyler Nading assisted with protocol selection. Additionally, Amy Pruden, Charles Bott, Chris Wilson, and all other previously mentions co-authors assisted with general guidance of experimental design and manuscript review.

Chapter 3: *Succession of the Resistome Through an Advanced Oxidation Process-Biologically Active Carbon Filtration-Based Treatment Train for Potable Reuse*

Chapter 3 primarily focuses on Research Objective (2) and assesses the impact of multiple stages of advanced water treatment on ARG profiles as determined by metagenomic sequencing. This manuscript assesses the impact of physical, biological, and chemical treatments on ARG removal or enrichment and to what extent each component of treatment provides a barrier to ARG dissemination. Consistent changes in ARG abundances allowed for the identification of core and discriminatory ARGs between treatments while differences between bulk ARGs and those of clinical relevance was also determined.

This manuscript has been prepared for submission to *Environmental Science & Technology.*

Attributions: I led sample collection, in lab molecular processing, data analysis, and manuscript writing. Additionally, Peter Vikesland, Amy Pruden, Charles Bott and all other previously mentioned co-authors assisted with general guidance of experimental design and manuscript review.

Chapter 4: *Advancing Metagenomics as a Tool for Profiling Microbial Functional Capacity: Insight into Biological and Metabolic Capabilities of an Ozone- Biologically Active Carbon Filtration Potable Reuse Treatment Train*

This manuscript focuses on Research Objective (3) and the application of NGS for functional annotation within potable reuse applications. Metagenomic sequencing results were used to assess the changes in bulk functional capacity between and throughout each stage of treatment. Further, discriminatorily abundance functional genes were identified, characterized, and analyzed especially as they related to BAF. In addition to comprehensive characterizations, functional genes associated with specific functions of interest (e.g. horizontal gene transfer, nitrogen cycling, biofilm formation, DNA repair) were analyzed at each stage of treatment.

This manuscript has been prepared for submission to *Environmental Science & Technology.*

Attributions: I led sample collection, in lab molecular processing, data analysis, and manuscript writing. Additionally, Amy Pruden, Charles Bott and all other previously mentioned co-authors assisted with general guidance of experimental design and manuscript review.

Chapter 5: *Relating Microbial Community Composition and Dynamics to Treatment Performance in an Advanced Oxidation Process-Biologically Active Carbon Filtration-Driven Potable Reuse Treatment Train*

This manuscript focuses on Research Objective (4) and characterizes the dynamics associated with microbial community composition throughout an advanced water treatment train via 16S rRNA gene amplicon sequencing. Specifically, changes in the bulk microbial community structure are evaluated between treatment process with variability within treatment processes contextualized to operational changes and water quality parameters. Selection and removal of specific taxa within each treatment process was identified and contextualized to functional niches known to those taxa. Through which, the ability for microbial community profiling to provide insights into functionality and resiliency of treatment was demonstrated. Finally, dynamics associated with microbial succession were measured during dynamic operation of the BAC filters and GAC contactors.

This manuscript has been prepared for submission to *Water Research*.

Attributions: I led sample collection, in lab molecular analysis, data analysis, and manuscript writing. Ramola Vaidya was primarily responsible for pilot operation under Germano Salazar Benites's direct management. Additionally, Amy Pruden, Charles Bott, and all other previously mentions co-authors assisted with general guidance of experimental design and manuscript review.

Chapter 6: *What is the Difference between Conventional Drinking Water, Potable Reuse Water, and Non-potable Reuse Water? A Core Microbiome Perspective*

This manuscript is focused on objective (5), and contextualizes the microbial community structure of multiple water uses (conventional drinking water, potable reuse, and non-potable reuse) to each other. Variability within individual water use's community structures were further related to differences in treatment, disinfection, climate, regionality, and other identified factors. Finally, core and discriminatory taxa were also identified within and between each water use. Ultimately, this research advances the ability for NGS technologies to be utilized as a water quality screening tool and expands our understanding of the differences between potable and non-potable water within the 'One Water' paradigm.

This manuscript has been prepared for submission to *Environmental Science & Technology.*

Attributions: Emily Garner and Pan Ji led the collection and laboratory processing of archival samples, with specific contributions outlined in their previously published work and within Chapter 6's methodology. I led the collection of supplemental samples and their subsequent processing. I was also responsible for combining the entire dataset, standardizing collected metadata, developing a post-sequencing trimming protocol to overcome sequencing differences, reanalyzing the entire sample set, and leading manuscript writing. Amy Pruden and all other previously mentions co-authors assisted with manuscript review.

Chapter 7: *Conclusions*

This chapter summarizes the key findings of this research and general contributions to the field of environmental engineering and water industry. Recommendations for future research are also provided.

REFERENCES

1. Gleick, P. H. *Water in crisis*. vol. 100 (New York: Oxford University Press, 1993).

2. Subramani, A. & Jacangelo, J. G. Emerging desalination technologies for water treatment: A critical review. *Water Res.* **75**, 164–187 (2015).

3. Subramani, A., Badruzzaman, M., Oppenheimer, J. & Jacangelo, J. G. Energy minimization strategies and renewable energy utilization for desalination: A review. *Water Res.* **45**, 1907–1920 (2011).

4. Döll, P. *et al.* Impact of water withdrawals from groundwater and surface water on continental water storage variations. *J. Geodyn.* **59**–**60**, 143–156 (2012).

5. Kenny, J. F. *et al. Estimated use of water in the United States in 2005*. (2009).

6. Megdal, S. B., Gerlak, A. K., Varady, R. G. & Huang, L. Y. Groundwater Governance in the United States: Common Priorities and Challenges. *Groundwater* **53**, 677–684 (2015).

7. Wada, Y. *et al.* Modeling global water use for the 21st century: The Water Futures and Solutions (WFaS) initiative and its approaches. *Geosci. Model Dev.* **9**, 175–222 (2016).

8. Boretti, A. & Rosa, L. Reassessing the projections of the World Water Development Report. *npj Clean Water 2019 21* **2**, 1–6 (2019).

9. World Water Assessment Programme (Nations Unies). World Water Development Report 2018 | UN-Water. *United Nations Educational, Scientific and Cultural Organization, New York, United States* https://www.unwater.org/publications/world-water-development-report-2018 (2018).

10. Mekonnen, M. M. & Hoekstra, A. Y. Sustainability: Four billion people facing severe water scarcity. *Sci. Adv.* **2**, (2016).

11. Usepa, U. *Guidelines for water reuse, 2012*. (2012).

12. Rauch-Williams, T., Madison, P. E., Marshall, R. & Davis, D. J. Baseline Data to Establish The Current Amount Of Resource Recovery from WRRFs Preparation of Baseline Data to ESTABLISH THE CURRENT AMOUNT OF RESOURCE RECOVERY. (2018).

13. Puyol, D. *et al.* Resource recovery from wastewater by biological technologies: Opportunities, challenges, and prospects. *Front. Microbiol.* **7**, 2106 (2017).

14. Guest, J. S. *et al.* A new planning and design paradigm to achieve sustainable resource recovery from wastewater. *Environ. Sci. Technol.* **43**, 6126–6130 (2009).

15. Batstone, D. J., Hülsen, T., Mehta, C. M. & Keller, J. Platforms for energy and nutrient recovery from domestic wastewater: A review. *Chemosphere* **140**, 2–11 (2015).

16. Matassa, S., Batstone, D. J., Hülsen, T., Schnoor, J. & Verstraete, W. Can direct conversion of used nitrogen to new feed and protein help feed the world? *Environ. Sci. Technol.* **49**, 5247–5254 (2015).

17. Daigger, G. T. Evolving Urban Water and Residuals Management Paradigms: Water Reclamation and Reuse, Decentralization, and Resource Recovery. *Water Environ. Res.* **81**, 809–823 (2009).

18. McCarty, P. L., Bae, J. & Kim, J. Domestic wastewater treatment as a net energy producer-can this be achieved? *Environ. Sci. Technol.* **45**, 7100–7106 (2011).

19. Tontti, T., Poutiainen, H. & Heinonen-Tanski, H. Efficiently Treated Sewage Sludge Supplemented with Nitrogen and Potassium Is a Good Fertilizer for Cereals. *L. Degrad. Dev.* **28**, 742–751 (2017).

20. Larsen, T. A., Alder, A. C., Eggen, R. I. L., Maurer, M. & Lienert, J. Source separation:

Will we see a paradigm shift in wastewater handling? *Environ. Sci. Technol.* **43**, 6121–6125 (2009).

21. Batstone, D. J. & Virdis, B. The role of anaerobic digestion in the emerging energy economy. *Curr. Opin. Biotechnol.* **27**, 142–149 (2014).

22. Shen, Y., Linville, J. L., Urgun-Demirtas, M., Mintz, M. M. & Snyder, S. W. An overview of biogas production and utilization at full-scale wastewater treatment plants (WWTPs) in the United States: Challenges and opportunities towards energy-neutral WWTPs. *Renew. Sustain. Energy Rev.* **50**, 346–362 (2015).

23. Daigger, G. T. Wastewater management in the 21st century. *J. Environ. Eng.* **133**, 671–680 (2007).

24. Smith, C. D. M. Potable reuse compendium. *Washington, DC USEPA* (2017).

25. Environmental Protection Agency, U. National Water Reuse Action Plan Draft. (2019).

26. Paulson, C., Broley, W. & Stephens, L. *Blueprint for One Water*. (2017).

27. Rice, J., Wutich, A. & Westerhoff, P. Assessment of de facto wastewater reuse across the U.S.: Trends between 1980 and 2008. *Environ. Sci. Technol.* **47**, 11099–11105 (2013).

28. Carey, R. O. & Migliaccio, K. W. Contribution of wastewater treatment plant effluents to nutrient dynamics in aquatic systems. *Environ. Manage.* **44**, 205–217 (2009).

29. Asano, T., Burton, F. & Leverenz, H. *Water Reuse: Issues, Technologies, and Applications*. *McGrawHill, New York,* (McGraw-Hill Education, 2007).

30. Leverenz, H. L., Tchobanoglous, G. & Asano, T. Direct potable reuse: A future imperative. *J. Water Reuse Desalin.* **1**, 2–10 (2011).

31. Hering, J. G., Waite, T. D., Luthy, R. G., Drewes, J. E. & Sedlak, D. L. A changing framework for urban water systems. *Environ. Sci. Technol.* **47**, 10721–10726 (2013).

32. Drewes, J. E. & Horstmeyer, N. Recent Developments in Potable Water Reuse. in 269–290 (Springer, Cham, 2015). doi:10.1007/698_2015_341.

33. Müller, J., Drewes, J. E. & Hübner, U. Sequential biofiltration – A novel approach for enhanced biological removal of trace organic chemicals from wastewater treatment plant effluent. *Water Res.* **127**, 127–138 (2017).

34. Drewes, J. E. & Khan, S. J. Water reuse for drinking water augmentation. *Water Qual. Treat.* 11–16 (2011).

35. Anderson, J. The environmental benefits of water recycling and reuse. *Water Supply* **3**, 1–10 (2003).

36. Kalkan, Ç., Yapsakli, K., Mertoglu, B., Tufan, D. & Saatci, A. Evaluation of Biological Activated Carbon (BAC) process in wastewater treatment secondary effluent for reclamation purposes. *Desalination* **265**, 266–273 (2011).

37. Benotti, M. J. *et al.* Pharmaceuticals and endocrine disrupting compounds in U.S. drinking water. *Environ. Sci. Technol.* **43**, 597–603 (2009).

38. Westerhoff, P., Yoon, Y., Snyder, S. & Wert, E. Fate of endocrine-disruptor, pharmaceutical, and personal care product chemicals during simulated drinking water treatment processes. *Environ. Sci. Technol.* **39**, 6649–6663 (2005).

39. Yoon, Y., Westerhoff, P., Snyder, S. A. & Wert, E. C. Nanofiltration and ultrafiltration of endocrine disrupting compounds, pharmaceuticals and personal care products. *J. Memb. Sci.* **270**, 88–100 (2006).

40. Yoon, Y., Westerhoff, P., Snyder, S. A., Wert, E. C. & Yoon, J. Removal of endocrine disrupting compounds and pharmaceuticals by nanofiltration and ultrafiltration membranes. *Desalination* **202**, 16–23 (2007).

41. Snyder, S. A. *et al.* Role of membranes and activated carbon in the removal of endocrine disruptors and pharmaceuticals. *Desalination* **202**, 156–181 (2007).

42. Ryu, J., Yoon, Y. & Oh, J. Occurrence of endocrine disrupting compounds and pharmaceuticals in 11 WWTPs in Seoul, Korea. *KSCE J. Civ. Eng.* **15**, 57–64 (2011).

43. Yoon, Y., Ryu, J., Oh, J., Choi, B. G. & Snyder, S. A. Occurrence of endocrine disrupting compounds, pharmaceuticals, and personal care products in the Han River (Seoul, South Korea). *Sci. Total Environ.* **408**, 636–643 (2010).

44. Ren, X., Chen, C., Nagatsu, M. & Wang, X. Carbon nanotubes as adsorbents in environmental pollution management: A review. *Chem. Eng. J.* **170**, 395–410 (2011).

45. Andersen, H., Siegrist, H., Halling-Sørensen, B. & Ternes, T. A. Fate of Estrogens in a Municipal Sewage Treatment Plant. *Environ. Sci. Technol.* **37**, 4021–4026 (2003).

46. Al-Hamadani, Y. A. J. *et al.* Sonocatalytical degradation enhancement for ibuprofen and sulfamethoxazole in the presence of glass beads and single-walled carbon nanotubes. *Ultrason. Sonochem.* **32**, 440–448 (2016).

47. Jung, C. *et al.* Adsorption of selected endocrine disrupting compounds and pharmaceuticals on activated biochars. *J. Hazard. Mater.* **263**, 702–710 (2013).

48. Han, J., Liu, Y., Singhal, N., Wang, L. & Gao, W. Comparative photocatalytic degradation of estrone in water by ZnO and TiO2 under artificial UVA and solar irradiation. *Chem. Eng. J.* **213**, 150–162 (2012).

49. Ryu, J., Oh, J., Snyder, S. A. & Yoon, Y. Determination of micropollutants in combined sewer overflows and their removal in a wastewater treatment plant (Seoul, South Korea). *Environ. Monit. Assess.* **186**, 3239–3251 (2014).

50. Kim, S. *et al.* Removal of contaminants of emerging concern by membranes in water and wastewater: A review. *Chem. Eng. J.* **335**, 896–914 (2018).

51. WateReuse. *California WateReuse Action Plan*. https://watereuse.org/wp-

content/uploads/2019/07/WateReuse-CA-Action-Plan_July-2019_r5-2.pdf (2019).

52. Olivieri, A. W., Pecson, B., Crook, J. & Hultquist, R. California water reuse—Past, present and future perspectives. *Adv. Chem. Pollution, Environ. Manag. Prot.* **5**, 65–111 (2020).

53. Bernados, B. Reverse Osmosis for Direct Potable Reuse in California. *J. Am. Water Works Assoc.* **110**, 28–36 (2018).

54. State Water Resources Control Board. Title 22 Code of Regulations. *Calif. Regul. Relat. to Drink. Water* (2019).

55. Van der Bruggen, B., Mänttäri, M. & Nyström, M. Drawbacks of applying nanofiltration and how to avoid them: A review. *Sep. Purif. Technol.* **63**, 251–263 (2008).

56. Matin, A., Laoui, T., Falath, W. & Farooque, M. Fouling control in reverse osmosis for water desalination & reuse: Current practices & emerging environment-friendly technologies. *Sci. Total Environ.* **765**, 142721 (2021).

57. Wenten, I. G. & Khoiruddin. Reverse osmosis applications: Prospect and challenges. *Desalination* **391**, 112–125 (2016).

58. Arola, K., Van der Bruggen, B., Mänttäri, M. & Kallioinen, M. Treatment options for nanofiltration and reverse osmosis concentrates from municipal wastewater treatment: A review. *https://doi.org/10.1080/10643389.2019.1594519* **49**, 2049–2116 (2019).

59. Malaeb, L. & Ayoub, G. M. Reverse osmosis technology for water treatment: State of the art review. *Desalination* **267**, 1–8 (2011).

60. Vaidya, R. *et al.* Pilot Plant Performance Comparing Carbon-Based and Membrane-Based Potable Reuse Schemes. *Environ. Eng. Sci.* **36**, 1369–1378 (2019).

61. Plumlee, M. H., Stanford, B. D., Debroux, J. F., Hopkins, D. C. & Snyder, S. A. Costs of Advanced Treatment in Water Reclamation. *http://dx.doi.org/10.1080/01919512.2014.921565* **36**, 485–495 (2014).

62. Steen, A. D. *et al.* High proportions of bacteria and archaea across most biomes remain uncultured. *ISME J. 2019 1312* **13**, 3126–3130 (2019).

63. Hug, L. A. *et al.* A new view of the tree of life. *Nat. Microbiol. 2016 15* **1**, 1–6 (2016).

64. Yarza, P. *et al.* Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol. 2014 129* **12**, 635–645 (2014).

65. Sivaganesan, M., Varma, M., Siefring, S. & Haugland, R. Quantification of plasmid DNA standards for U.S. EPA fecal indicator bacteria qPCR methods by droplet digital PCR analysis. *J. Microbiol. Methods* **152**, 135–142 (2018).

66. Shanks, O. C. *et al.* Data Acceptance Criteria for Standardized Human-Associated Fecal Source Identification Quantitative Real-Time PCR Methods. *Appl. Environ. Microbiol.* **82**, 2773–2782 (2016).

67. Shokralla, S., Spall, J. L., Gibson, J. F. & Hajibabaei, M. Next-generation sequencing technologies for environmental DNA research. *Mol. Ecol.* **21**, 1794–1805 (2012).

CHAPTER 2: BIODEGRADABLE DISSOLVED ORGANIC CARBON PROFILING REVEALS CAPACITY OF CARBON-BASED POTABLE REUSE TREATMENT

Matthew F. Blair, Ramola Vaidya, Germano Salazar Benites, Larry Schimmoller, Tyler Nading, Chris Wilson, Amy Pruden, and Charles Bott

ABSTRACT

Biological treatment is gaining ground as a means to enhance removal of total organic carbon (TOC) as part of a multi-barrier treatment train for water reuse. Here we applied biodegradable organic carbon (BDOC) analysis to evaluate the extent of removal of various TOC fractions through a pilot-scale water reuse train employing flocculation/sedimentation, ozone, biologically-active carbon (BAC), and granular activated carbon (GAC). BDOC analysis highlighted GAC and ozone treatments as critical to non-biodegradable dissolved organic carbon removal and the need to optimize BAC performance to maximize GAC adsorption capacity. BDOC analysis was further applied to benchmark process performance to operational conditions, such as empty bed contact time, occurrence of nitrification, and operational upsets. Overall, BDOC analysis proved an asset for understanding and improving operation of ozone/BAC/GAC treatments for water reuse.

INTRODUCTION

Advanced water treatment (AWT) for the purpose of producing high quality water for potable reuse can take many forms, typically employing multiple physical and chemical treatment barriers. Reverse osmosis is often thought of as an essential treatment barrier for potable reuse [1,2], given its ability to remove a wide range of contaminants. However, reverse osmosis presents several drawbacks, including high energy demands, loss of efficiency due to fouling, high implementation costs, and need to treat and dispose of the brine waste stream [3–5]. Further, the quality of water may not be compatible with the aquifer geochemistry for recharge. This has brought about interest in membrane-free, AWT trains as an attractive alternative. In particular, carbon-based AWT trains employ ozone and biologically active carbon (BAC) coupled with granular activated carbon (GAC). This alternative harnesses the natural biodegradative capacity of microorganisms inhabiting the BAC to provide biological degradation of trace organic contaminants, such as pharmaceuticals and personal care products, which are inefficiently removed in conventional drinking water and wastewater treatment trains [6–8]. To expand the application of ozone-BAC-GAC, it is necessary to gain a better understanding of its capacity to remove various forms of organic carbon; including total organic carbon (TOC), dissolved organic carbon (DOC), biodegradable dissolved organic carbon (BDOC), non-biodegradable dissolved organic carbon (NBDOC), and particulate organic carbon (POC), in the face of various long-term operational changes and upsets.

BAC filtration has become widely applied in the context of conventional drinking water treatment for removing organic carbon [9–13], including precursors to disinfection byproduct formation [14]. The capacity of BAC filtration can be further enhanced by employing an upstream advanced-oxidation process, such as ozone, which acts to increase the bioavailability of complex organic compounds ahead of the biofilter [15]. However, the knowledge base of BAC filtration for conventional drinking water treatment does not necessarily directly translate to its optimization

for potable reuse. For example, the organic matter makeup in a wastewater's tertiary effluent consists largely of soluble microbial products and is distinct from the natural organic matter encountered in more conventional drinking water sources [16,17]. Further, a key consideration for the ozone-BAC-GAC configuration is to maximize biodegradation of DOC in the BAC to the extent possible to maximize the lifespan of the GAC, where BDOC and (much more difficult to remove) NBDOC will equally exhaust GAC sorptive capacity, leading to the costly need to regenerate GAC. The effectiveness of biofiltration has been previously linked to changes in empty bed contact time (EBCT), ozone dose, temperature (and related kinetics), nutrient and substrate matrixes, and lesser understood microbial interactions [9,18–21]. Treatment train performance can be further improved through the implementation of upstream, synergistic processes such as coagulation-flocculation-sedimentation (Floc-Sed), which can decrease contaminant loads to subsequent ozonation and biofiltration by removing BDOC, NBDOC, and POC [22–24].

Bench-scale testing and analysis protocols could help to better understand and optimize ozone-BAC-GAC for potable reuse applications. Ideally, bench testing would provide a straightforward and cost-effective means to assess treatment efficacy. TOC is often monitored as a gross indicator of water quality through a treatment train, as it is typically cost prohibitive to monitor each individual trace organic contaminant of concern [25]. However, to optimize biological treatment, it is essential to be able to assess the biodegradable TOC fraction. For example, approximately 80% of conventional TOC encountered in surface water sourced drinking water [10] and secondary wastewater effluent [26] is non-biodegradable.

Various means of assessing BDOC have been applied to characterizing potable water [27–30]. The assimilable organic carbon test measures cellular growth of two model strains of bacteria that commonly occur in drinking water, relating it to measured loss of TOC and yielding units of mg acetate equivalents/L [31,32]. However, a drawback of this approach is that it is unlikely that the two strains employed will be capable of degrading the full spectrum of biodegradable constituents represented by a given TOC measurement. The BDOC test circumvents this problem by employing an inoculum of microorganisms that is native to the system being tested and thus is more likely to yield a representative assessment of the true proportion of TOC that is biodegradable. Conventionally, the BDOC test is applied in an analogous fashion as the biochemical oxygen demand test, where TOC (rather than dissolved oxygen) is monitored in the bulk water with time. However, a solid media-based approach was recently developed that is more relevant to biofilm-based treatments, such as BAC filtration [33–35]. When applied across an AWT train, the BDOC test could ideally serve to inform with respect to the percent biodegradable carbon removed by each treatment stage, providing a direct measurement of the theoretical maximum possible DOC removal via biodegradation, the nonbiodegradable fraction, and overall capacity of biological treatment. Here, BDOC analysis was used to assess the capacity for removal of TOC, DOC, BDOC, NBDOC, and POC through a pilot-scale carbon-based AWT potable reuse treatment train employing coagulation/flocculation/sedimentation (Floc-Sed), ozone, BAC, GAC, and UV disinfection. Specifically, the objectives of this research were to: 1) evaluate how each treatment stage influences overall TOC removal and the proportion of POC, BDOC, and NBDOC removed and 2) characterize how changes in operating conditions (e.g., ozone dose, monochloramine addition, nutrient addition, and temperature) and upset events (i.e., individual treatment processes operating outside of intended performance

range or uncharacteristic changes in water quality parameters) are reflected in BDOC measurements and overall performance over long-term operation. Of particular interest to this study was the effect of EBCT on organic carbon profiles following BAC filtration and GAC contacting.

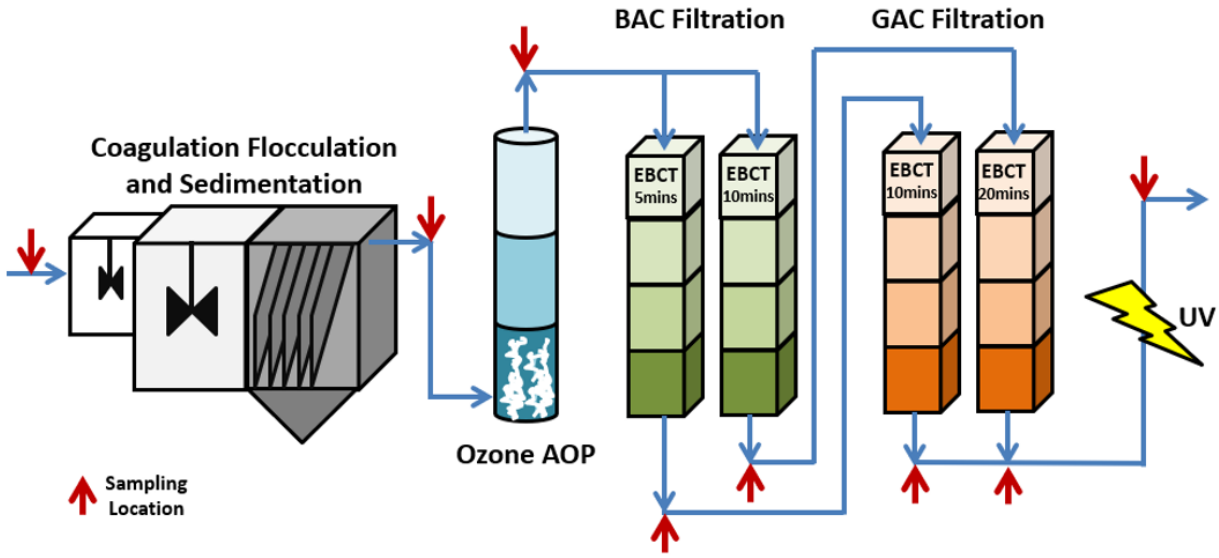


Figure 2.1: Hampton Roads Sanitation District SWIFT Pilot Facility's process schematic. Sampling locations indicated with red arrows. The pilot facility was fed tertiary treated denitrified effluent from the York River wastewater treatment plant at a flowrate of 4.3 gallons per minute.

METHODS

Site description, sample collection, and preservation

Samples were collected from a 4.3 gpm AWT pilot plant operated at the Hampton Roads Sanitation District's York River Wastewater Treatment Plant (Figure 2.1) over an ~18-month period. The pilot-scale plant uses York River's tertiary effluent (secondary effluent treated via a methanol supplemented dentification filter - TETRA™ DeepBed™ filters (De Nora, Milan, Italy)) and employs coagulation-flocculation-sedimentation (Floc-Sed), monochloramine supplemented ozonation, BAC filtration, GAC contacting, and ultraviolet disinfection. To assess the impact of EBCT, the pilot flow was split between two different BAC-GAC trains: one with 5 min BAC EBCT (BAC5) followed by 10 min GAC EBCT (GAC10) and the other with 10 min BAC EBCT (BAC10) followed by 20 min GAC EBCT (GAC20). The BAC effluent was not combined; the 10 min GAC EBCT unit received flow from only the 5 min BAC EBCT unit. In addition to the treatment processes themselves, various chemicals and operational conditions were adjusted periodically. Major changes and upset conditions are outlined in Table 2.1 with additional design parameters provided in the SI (SI Table A.1-A.2 and SI Figure A.1-2). Prior to this period of testing, the pilot was operated under relatively stable conditions from June 2016 until beginning the addition of preformed monochloramine in September 2016, substantially before the initial sampling in January of 2017. Further, BAC filters were exhausted prior to the pilot commissioning through extended operation with secondary effluent until organics removal plateaued. Sampling events were also conducted at intervals ≥ 3 weeks following process

modification, with most occurring ~7 weeks. Operating conditions were stable according to standard water quality parameter monitoring, unless otherwise noted. Additional details on pilot operation during this study period are available in prior publications [25,36–39] with summarized water quality data for the pilot influent and GAC effluent provided in SI Table A.1**.**

Bulk water samples for BDOC testing were collected at the effluent of each treatment, aligning hydraulic residence time to the extent possible to ideally capture the same plug of water as it traveled through the pilot. Sampling events and dates are reported in Table 2.1**.** For each sample, two 1-liter samples were collected in acid-washed, baked glass bottles with triplicate 25-mL samples collected in acid-washed, baked 40-mL glass vials for both TOC and DOC analysis. DOC samples were syringe-filtered using a sterile PVDF 0.45-µm syringe filter (MilliporeSigma™ Millex™, Darmstadt, Germany), which was rinsed immediately before collection with sample water to prevent carbon leaching from the filter membrane itself. Both TOC and DOC samples were acid preserved (pH<2) using 85% phosphoric acid before they, and the non-acid preserved 2-liter samples, were stored on ice for transit.

BDOC Analysis Preparation and Setup

BDOC analysis was conducted in accordance with previously published methods [33,34] with minor alterations to samples volumes and inoculum preparation. At-least one week prior to sample collection, a biologically-active sand column (BAS) was set up and seeded with conditioning water, a onetime seed slurry of HACH polyseed and nutrients, and then continuously fed pilot influent water to develop an attached microbial community for inoculation into the BDOC testing bottles. The initial sand (effective size ranging from 0.45mm to 0.55mm with a uniformity coefficient of 1.5) was rinsed with DI water and baked at 550 C° for 5.5 hours to burn off residual carbon.

After collection, water samples and the BAS reactor were transported to the laboratory where the samples and BAS were combined. Initially, the BAS was gently rinsed three times with deionized water to remove residual organic carbon from the feeding phase and prevent contamination of the BDOC samples, especially those with lower DOC concentrations. After the first wash, ~50 grams BAS (wet weight) was inoculated into acid washed and baked BDOC bottles (triplicate 250mL-bottles for each sample location). Following inoculation, each triplicate set of BDOC bottles and inoculated sand was rinsed three times using the first liter of sample water, to further prevent contamination from the inoculated BAS which was originally conditioned with pilot feed water. After the second washing, each bottle was filled with approximately 200 mL of its corresponding sample, gently shaken, and sampled for a reactor DOC measurement. All sample bottles were placed on a shaker table set at 100 rpm for approximately 30 days at 20±2 C° and sampled for DOC measurements every ~7 days.

Laboratory TOC and DOC Measurements

All TOC/DOC acid preserved samples were analyzed using a Sievers 5310 C Laboratory TOC Analyzer (Suez, Boulder, CO), which was purged prior to operation to prevent run to run contamination. Prior to analysis, each sample was sparged with nitrogen gas for 3 minutes at 10 PSI to remove inorganic carbon. Standards and nanopure water blanks were run at the beginning, end, and every 10-12 samples. Each sample was measured three times, with the first

measurement being considered a “rinse” sample and the remaining two TOC and DOC measurements averaged to represent the measurement for each biological replicate.

Water Quality and Operational Conditions Data Collection

BDOC measurements were compared to operational and water quality data. This comparison leverages periodic water quality monitoring occurring at the SWIFT facility, utilizing a variety of standard methods and online operational data collected continuously during treatment operation with inline sensors.

Because water quality data were collected periodically, sampling dates were used to sort and identify data collected around each sampling event, specifically the day of and the three days preceding and following each sampling date. Because all parameters were not measured every day of operation, selection criteria were developed that prioritized: 1) data collected on the day of sampling, 2) average of the day before and after sampling, 3) the only value in the data range if only one existed, or 4) a blank value. After selection, the data were used to benchmark BDOC measurements to water quality characteristics.

General operational data and parameters were collected by online analyzers at approximately 5-minute intervals. Daily summary statistics, including averages, standard deviations, minima, and maxima were calculated and used to benchmark treatment operational data to subsequent BDOC measurements.

Data Analysis and Statistics

DOC measurements over the ~30-day BDOC test were plotted to produce carbon degradation curves (e.g., SI Figure B.1)**.** The initial and final DOC measurements were subtracted to calculate the amount of organic carbon remaining after treatment that was biodegradable (BDOC, mg/L). The ratio of the BDOC to the initial DOC measurement was calculated as the biodegradable fraction, with the final DOC measurement indicating the amount of NBDOC. The change in initial DOC measurements (i.e., effluent DOC from each treatment stage) between treatment processes was used to calculate the bulk DOC removal achieved in the pilot (mg/L) and the total treatment efficiency (% removal) by taking the ratio of DOC removed to initial DOC. Further, initial TOC and DOC measurements were used to calculate the amount and fraction of carbon that was found in the dissolved and particulate phases. Combinations of the above were utilized to fractionalize the carbon profile through the AWT train and assess carbon removal performance.

Non-parametric Wilcoxon signed-rank tests were conducted in R version 4.0.2 (Rstudio Team, 2020) to compare changes in organic carbon components between sampling location. Following the application of Wilcoxon signed rank tests on all samples, outliers were removed when justified by cross-referencing to sampling notes and the identification of documented upset conditions, prior to conducting exploratory correlation analysis. The distribution of each variable was tested for normality utilizing the Shapiro-Wilk test in Rstudio. Variables with non-normal distributions were removed from analysis and the remaining normally-distributed variables were correlated to water quality and operational data using the Pearson linear correlations. To correct for false discoveries associated with multiple comparisons, the Benjamini-Hochberg (BH) procedure was applied to each set of correlation tests at a liberal FDR set at 0.20, per exploratory

recommendations [41]. All data were also plotted (Appendix B Subsection 2) to assess temporal variation. Significance for all statistical tests was set at $p<0.05$.

RESULTS AND DISCUSSION

Operational Conditions and Organic Carbon Profiling through the Carbon-Based AWT Train

Table 2.1 provides a timeline of the ten sampling dates examined in this study as they relate to operational conditions and upset events. Each sampling event is named sequentially (S01-S10) and further differentiated according to whether the event represented stable (e.g., SS01) or upset (e.g., SU02) conditions or any other relevant defining information (e.g., SUMeth02, SUTurbid09). Monochloramine addition (3 mg/L Cl_2) and ozone ratios were sequentially adjusted, to suppress formation of bromate, but required dechlorinating with sodium bisulfite or calcium thiosulfate to prevent disinfectant from entering microbially-active BAC filters. Following months of stable operation targeting 1:1 ozone:TOC dose, the ozone ratio was decreased to assess its impact on BDOC generation. Sampling events following SS06 sequentially targeted ozone:TOC ratios of 0.5, 0.25, 1 and 0.8. Outside of the intentional operational variation, performance was also likely affected by upsets of increased DOC loading during SUMeth02 and SUMeth04, due to carry-over of methanol from the upstream denitrification unit, and during SUTurbid09, due to a Floc-Sed failure and elevated turbidity. Also, ozone dose variability (i.e., standard deviation) decreased from SS05 onwards, when more precise control measures were implemented. GAC performance indicated dynamic shifts in carbon removal with time, consistent with adsorption sites being exhausted and shift to biological treatment as a function of bed volumes passed through the filter (Appendix B Subsection 1). All but one sample was collected after 18,000 bed volumes, at which point the GAC would be expected to be largely exhausted. SI Table B.1 summarizes the observed organic carbon profiles for each sampling date, including TOC, DOC, BDOC, NBDOC, and POC.

Table 2.1: Overview of Sampling Events, Operational Conditions, and Upset Events

Sampling Event	Date Sampled or Started	Specific Operational Conditions and Changes	System Upsets and Anomalies
-	Before 09/01/2016	**Chlorinated Backwashed Start (0.2 mg/L Cl_2)**	-
-	09/01/2016	**Preformed Monochloramine Dosing Start (3 mg/L Cl_2)**	-
SSInital-01	01/19/2017	-	Pilot influent not sampled
SUMeOH-02	02/01/2017	-	Methanol spike in pilot influent from upstream denitrification
SSInfSwitch-03	02/16/2017	Fed pilot with a different plant's secondary effluent*	
-	02/22/2017	**Chlorine Quenching Started Prior to BAC (Bisulfite)**	-
-	03/20/2017	**Chlorinated Backwash Stopped**	-

SUMeOH-04	03/30/2017	-	Methanol spike in pilot influent from upstream denitrification
-	04/01/2017	-	Media Loss (16 inches of BAC5 replaced with 8 inches from BAC10 with adjusted flow to maintain EBCT)
SS-05	06/22/2017	-	-
SS-06	08/22/2017	-	-
-	09/18/2017	**Nutrient Additions Start** Ammonia $(NH_4)_2SO_4$ 2 mg/L as N Phosphorous H_3PO_4 0.2 mg/L as P	-
-	10/10/2017	**Pilot influent temperature decreased using chiller(15° C)**	-
SSChillerNutrients-07	12/12/2017	0.5:1 O3:TOC, Chilled (15°C), Calcium Thiosulfate Quench, Additional Nutrients	S2 initial TOC|DOC measurement error
SSLowOzone-08	01/31/2018	0.25:1 O3:TOC, Chilled (15°C), Calcium Thiosulfate Quench, Additional Nutrients (N, P)	-
-	-	-	Media Loss (5 inches of BAC5 lost with adjusted flow to maintain EBCT)
-	-	**Filter Aid Polymer Added**	-
SUPolymerTurbid-09	02/14/2018	1:1 O3:TOC, Chilled (15°C), Bisulfite Quench, Additional Nutrients (N, P), Filter Polymer	Low coagulant dosing (stock issues) led to issues with FlocSed performance and elevated turbidity.
-	03/6/2018	**Chlorinated Backwashes Start (0.2 mg/L Cl_2)**	-
-	-	**Residual Chlorine Loading onto the BAC filters without Quenching (3 mg/L Cl_2)**	-
SSNoQuench-10	04/02/2018	0.8:1 O3:TOC, Chilled (15°C), No Quench, Additional Nutrients (N, P), Filter Polymer	-

*Not included in analysis

Fate of DOC in the Carbon-Based AWT Train

Performance of the carbon-based AWT train was evaluated across all ten sampling dates according to the conventional measure of DOC removal (Figure 2.2) (SI Table A.1). Floc-Sed significantly reduced DOC (Wilcoxon: $p<0.05$), with a median reduction of 1.65 mg/L (24% of the influent carbon). Ozonation did not significantly reduce DOC, with a median change of -0.02 mg/L (SI Table A.2), but did result in an average UVT increase of 8.80% ± 2.55 as a result of breakdown of organic compounds. Both BAC5 and BAC10 provided reductions of DOC (Wilcoxon: $p<0.05$), with median removals of 0.65 mg/L (11% treatment efficiency) and 1.19 mg/L (19% treatment efficiency), respectively. The observed improvement in DOC removal with elevated EBCT was significant (Wilcoxon: $p<0.05$). Similarly to the BAC filters, both GAC10 and GAC20 provided reductions in DOC (Wilcoxon: $p<0.05$), with median reductions of 0.93 mg/L (19% treatment efficiency) and 1.04 mg/L (27% treatment efficiency), respectively. In contrast to BAC filtration, increased EBCT did not result in additional reduction of DOC concentrations by GAC treatment. However, the BAC10/GAC20 combination did provide improved DOC removal relative to the BAC5/GAC10 combination (Wilcoxon: $p<0.05$).

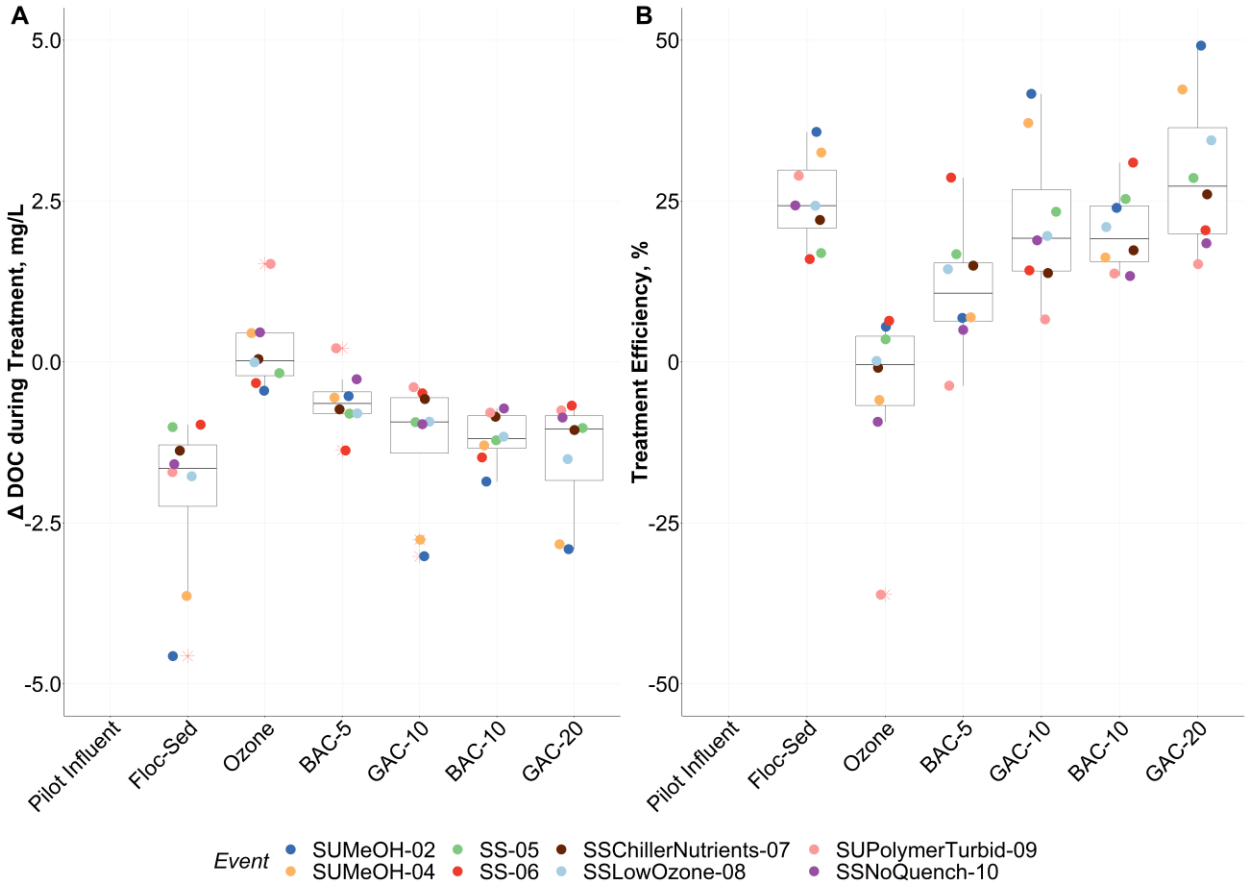


Figure 2.2: Dissolved organic carbon (DOC) removal of each treatment stage relative to the prior treatment stage for all ten sampling events: A) mass change of DOC (mg/L) and B) percent DOC removal relative to DOC measured in prior treatment stage effluent.

Under the conditions of this study, increased EBCT resulted in increased DOC removed during BAC filtration. Increased contact time likely increased interactions between microorganisms and BDOC and facilitated biological degradation of organics with slower kinetics [42,43]. Additionally, increased EBCT may enhance break-down of microbial inhibitors, such as ozone and associated by-products, effectively reducing their impact on the microbes. Notably, DOC removal did not increase with increased EBCT in the GAC contactors, as both GAC10 and GAC20 yielded similar net reductions in DOC concentrations. However, since upstream BAC performance was improved by EBCT, this resulted in a greater percent removal of the remaining DOC during GAC contacting. This highlights the importance of optimizing upstream BAC filtration to maximize GAC contacting and prolong its operational lifespan. Still, the remaining organic carbon fraction was 29% and 22% BDOC, respectively, indicating that there is still potential for biofouling and microbial regrowth in the final water [28,30,44–46].

Fate of POC in the Carbon-Based AWT Train

POC typically provides a non-negligible contribution to the overarching carbon profile and can be problematic in downstream distribution systems [47] leading to sedimentation, biofouling, clogging, microbial regrowth, and other water quality issues [48–54]. Therefore, assessing changes in POC is also useful to better understand potential downstream impacts of POC remaining from upstream WWTP processes, and biomass derived from the BAC processes.

Given that the influent water had already been subjected to conventional wastewater treatment with tertiary denitrification, POC levels in the feed were very low. However, it is important to consider that BAC filtration will convert DOC into biomass on the filter, which will elevate downstream POC proportionally to any sloughing/release of biomass into the bulk water. In this treatment train, though, POC was not notably elevated with biological treatment. Less than 0.5 mg/L of POC was measured across all sampling events and locations (SI Figure A.1), or 10% of the TOC loading (Figure 2.3a), with most samples < 5% POC. The average change in POC throughout each process only increased (indicated by the median measure of central tendency) following GAC treatment (Figure 2.3b), indicating that biomass was generated and released within the GAC units, but only after being greatly reduced relative to the influent. It is important to note, however, that the increase in POC following GAC contacting was small (average < 3% for GAC20, and < 0.1% for GAC10, SI Table A.1) and the resulting POC concentrations were also low (average 0.13 mg/L, SI Table A.2), especially when compared to the average influent POC concentration (0.55 mg/L, SI Table A.3). Further, POC following BAC filtration and/or GAC contacting may also include deteriorated activated carbon.

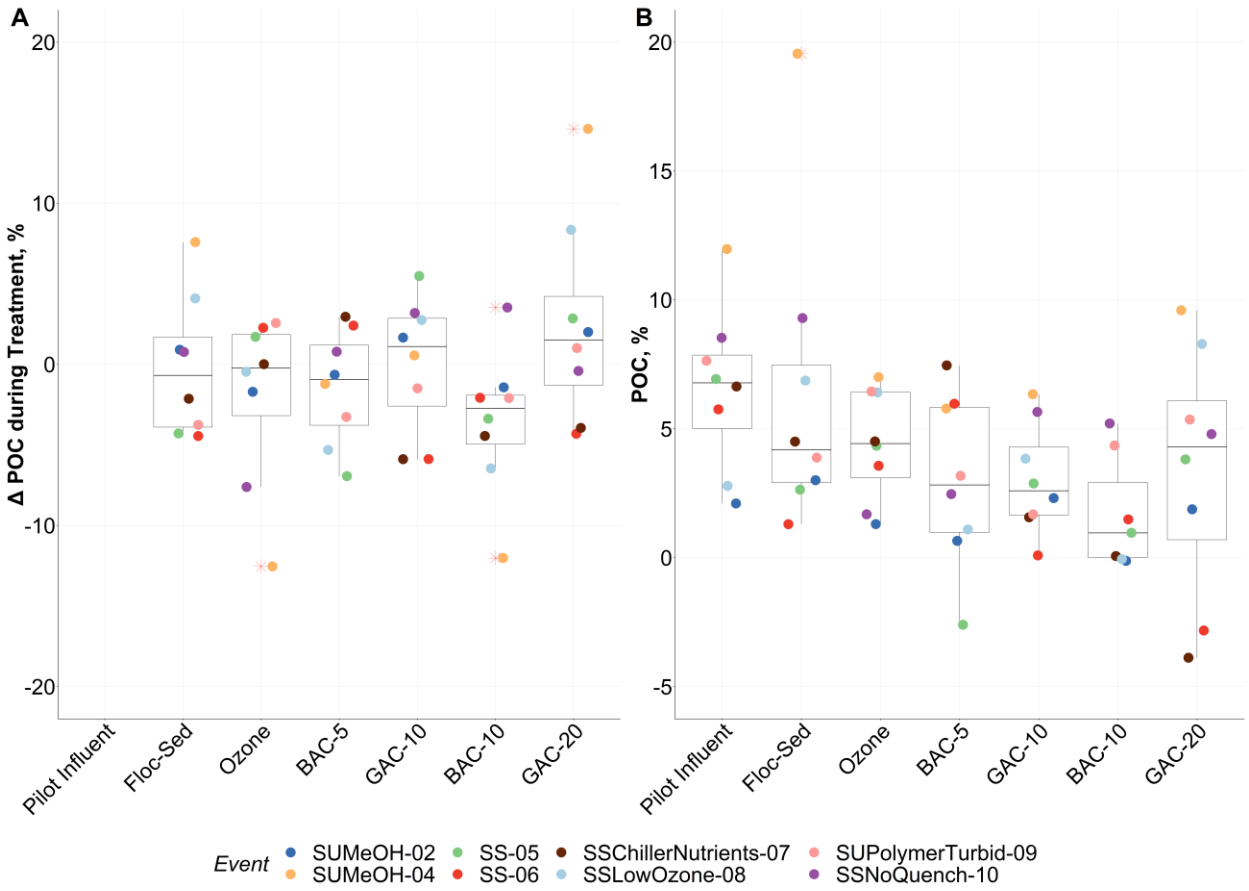


Figure 2.3: Particulate organic carbon (POC) measured at each stage of the treatment train: A) percentage (%) of POC relative to the total organic carbon (TOC) present and B) change in the percent (%) POC relative to the effluent POC of the prior state of treatment.

Although influent POC was low in this study, it is important to note that alternative configurations could result in much higher concentrations of POC entering the ozone unit [34,55–58]. An interesting finding in the present study was that POC largely appeared to be in a biodegradable form, but because the majority of POC first entered the ozonation unit, it primarily acted as an ozone sink. Minimal POC was produced from DOC being converted into biomass on the BAC/GAC contactors and the net levels of POC were negligible in the finished water compared to the influent.

Fate of BDOC in the Carbon-Based AWT Train

BDOC analysis provided deeper insight into the composition of the TOC through the treatment train, the role of each treatment stage, and performance with time. Figure 2.4 provides an example (Sampling Event SUMeth02 and SS-05) of the relative proportions of DOC that were BDOC and NBDOC through the treatment train and how much of each was removed by each stage. SI Table A.3 summarizes statistical comparisons. Figure 2.5 provides a more detailed accounting of BDOC concentrations and their dynamics throughout the AWT train across the ten sampling dates.

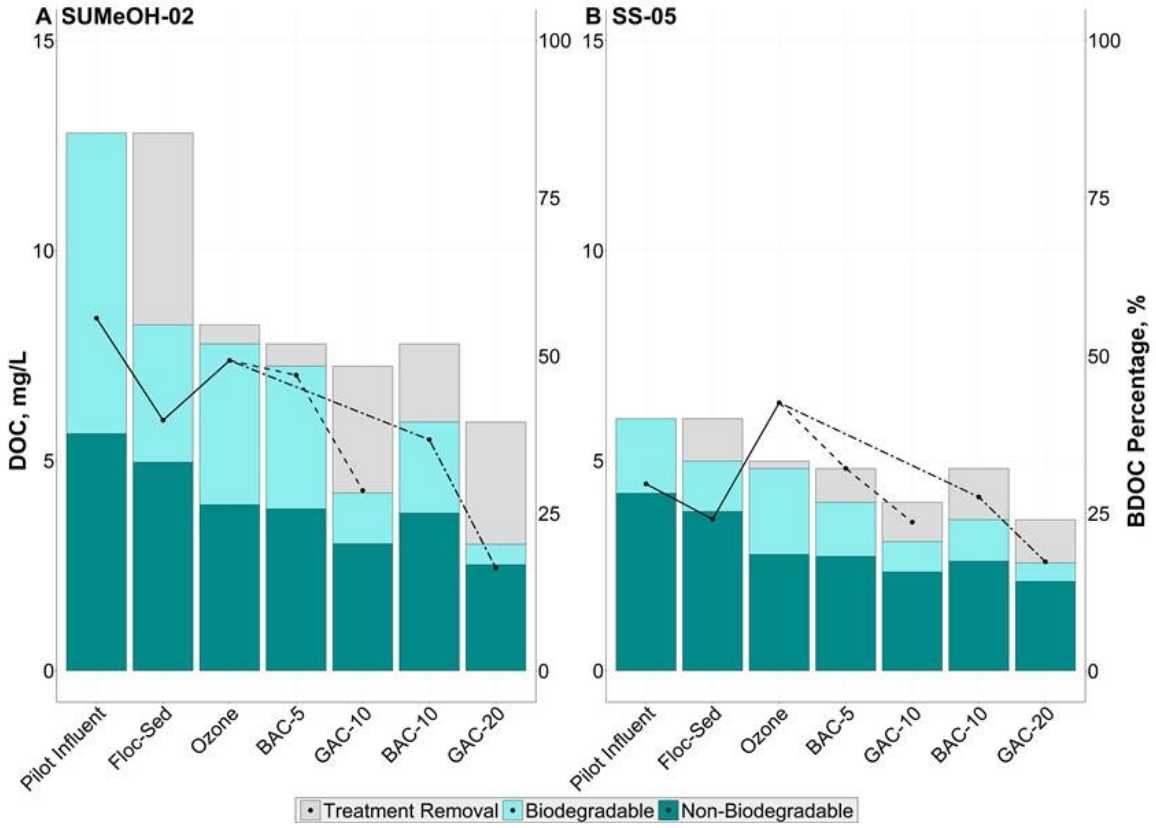


Figure 2.4: Dissolved organic carbon (DOC) measured in the effluent of each indicated treatment stage along the treatment train that is Non-Biodegradable or Biodegradable. The gray shaded area, "Treatment Removal," indicates the portion of DOC removed relative to the prior treatment stage. The line plot corresponds to the secondary y-axis and indicates the Percent Biodegradable DOC in the effluent of the indicated treatment stage. Exemplar sampling events include (A) SUMeOH-02, representative an upset condition impacted by upstream methanol leaching and (B) SS-05, representative of stable operation.

Aside from two of the upset sample events (SUMeth02, SUMeth04), pilot influent BDOC concentrations were relatively stable, ranging from 1.75 mg/L to 3 mg/L, representing 25% to 40% of the total DOC concentration (Figure 2.5e). Floc-Sed effectively decreased BDOC (Wilcoxon: $p<0.05$), with an average removal of 1.78 mg/L (SI Table A.1) and average percent removal of BDOC greater than 60%. Interestingly, the BDOC fraction was much more effectively reduced by Floc-Sed than the NBDOC fraction. This suggests that the particulates removed by Floc-Sed, such as micro-flocs lingering after upstream biological treatment, were largely biodegradable in nature.

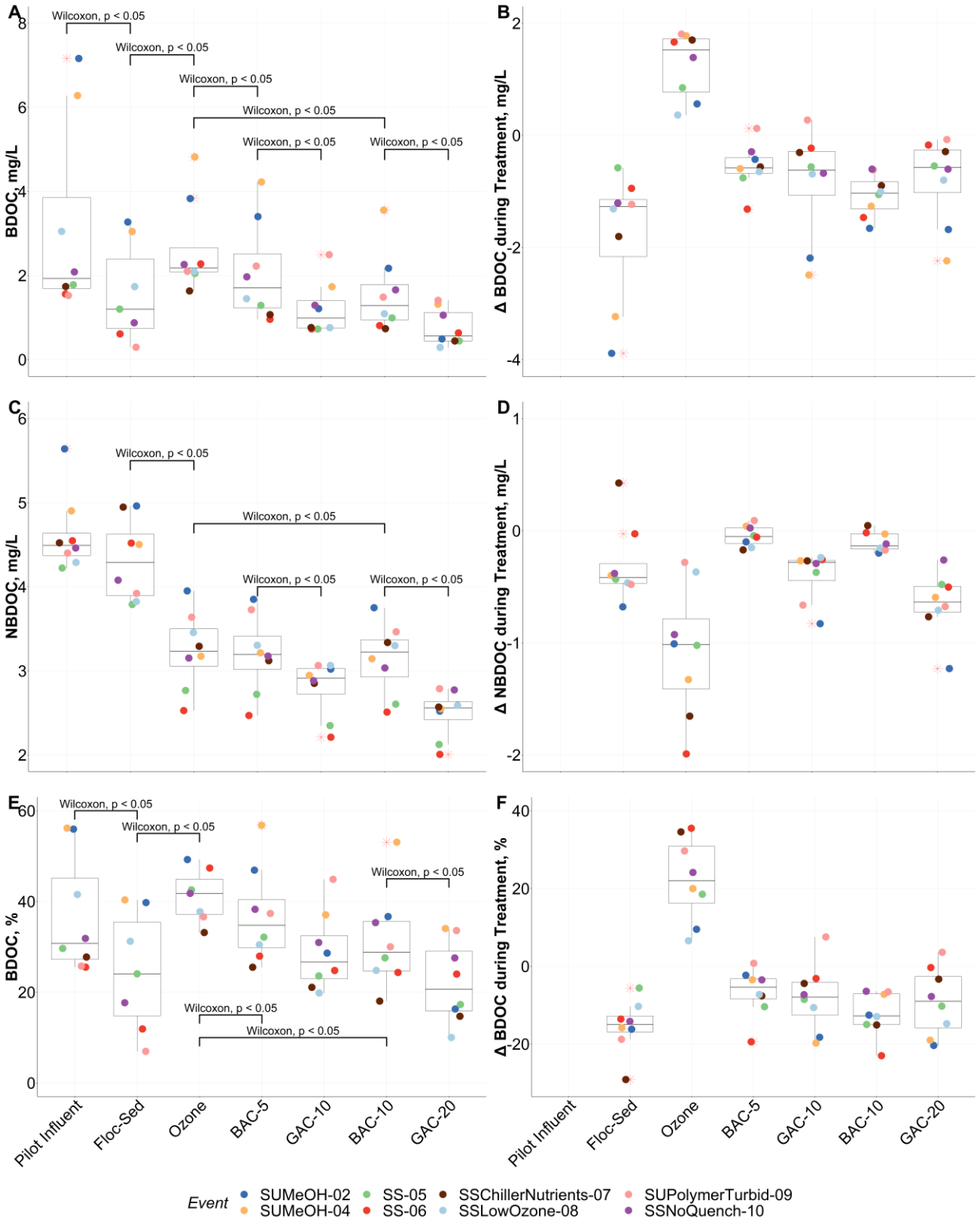


Figure 2.5: Biodegradable organic carbon (BDOC) and non-biodegradable organic carbon (NBDOC) measurements at each treatment stage: A) BDOC concentration (mg/L), B) change in BDOC concentration (mg/L) relative to prior stage of treatment, C) NBDOC concentration

(mg/L), D) change in NBDOC concentration (mg/L) relative to the prior stage of treatment, E) percentage (%) of BDOC, relative to DOC, remaining after treatment, and F) change in the percentage (%) of BDOC achieved by each stage of treatment.

Ozonation was the only the process that resulted in a significant increase in BDOC concentrations (Wilcoxon: $p<0.05$) and percent BDOC relative to total DOC (Wilcoxon: $p<0.05$). This indicates that ozonation successfully served its purpose of increasing the biodegradability of incoming organic matter, a result that could not be discerned from DOC analysis alone. Accordingly, both BAC5 and BAC10 achieved substantial reduction in BDOC (Wilcoxon: $p<0.05$) with BAC10, on average, yielding more extensive DOC removal (SI Table A.2). The GAC treatment stage provided further reduction in BDOC, with GAC10 averaging 0.86 mg/L of BDOC removed (Wilcoxon: $p<0.05$) and GAC20 averaging 0.80 mg/L of BDOC removed (Wilcoxon: $p<0.05$).

Fate of NBDOC in the Carbon-Based AWT Train

As summarized in Figure 2.4 and Figure 2.5, the recalcitrance of NBDOC through the carbon-based AWT train was apparent. Only ozonation, GAC10, and GAC20 yielded any measurable reduction in NBDOC (Wilcoxon: $p<0.05$) with average reductions of 1.07, 0.40, and 0.65 mg/L, respectively (SI Table A.1). Notably, ozonation did not provide measurable reduction in DOC, but the BDOC test revealed its ability to transform NBDOC into BDOC. Interestingly, although Floc-Sed substantially reduced BDOC, it failed to significantly change the NBDOC concentration. This suggests that NBDOC entering the pilot was largely in a soluble form and not amenable to charge neutralization and/or sorption to the sedimented flocs that formed.

In order to be able to discern removal mechanisms of biodegradation versus sorption through the treatment train, the GAC implemented in the BAC filters was intentionally exhausted prior to commencing this study, resulting in theoretically negligible sorption-based removal. Correspondingly, the BAC units did not substantially reduce NBDOC and the NBDOC concentrations in their respective effluents were equivalent to each other and the effluent from upstream ozonation. However, the BAC10 effluent NBDOC concentrations were significantly lower than that of the ozonation effluent (Wilcoxon: $p<0.05$), even if the achieved reductions were minimal. Ozone/BAC resulted in average NBDOC reduction of 1.12 mg/L and 1.18 mg/L for the low and high EBCT trains, respectively, with the ozonation process accounting for 1.07 mg/L.

Outside of significant changes to EBCT, further reductions of NBDOC were found to rely on optimizing the ozonation process and subsequent BAC filtration, resulting in decreased BDOC loading onto the GAC contactors and maximizing availability of adsorption sites for NBDOC. This important interplay between BAC and GAC performance, especially as it relates to NBDOC removal, was highlighted by the removal dynamics observed throughout the pilot treatment system. Of the total carbon removed during GAC treatment, 68.3% was biodegradable at the lower EBCT (GAC-10) while that percentage was decreased to 55.1% at the higher EBCT (GAC-20) with improved BDOC removal in the preceding BAC-10 filter. Considering that total DOC removal throughout both GAC contactors was fairly consistent, the increased reduction of BDOC during BAC filtration allowed GAC-20 to more effectively remove NBDOC. Theoretically, the complete removal of BDOC prior to GAC, i.e., through optimized BAC filtration, should result in the lowest effluent DOC concentration and greatest reductions in

NBDOC. This will extend the lifespan of the GAC, which, as time goes on, will lose adsorptive capacity and eventually functionally transition into BAC filters and need to be regenerated or replaced [59]. In the present study, the GAC units were operating primarily beyond 20,000 bed volumes, at which point it is generally understood that adsorptive capacity is largely spent. Yet, the continued NBDOC removal by GAC contacting revealed through the BDOC analysis in this study suggests that substantial adsorption was still occurring.

Correspondence of Organic Carbon Profiling to Water Quality Data and Ability to Assess Operational Changes and Diagnose Upset Events

The above trends and analyses were integrated across nearly 18 months of operation, including various operational changes and upsets (Table 2.1), indicating that overall organic carbon removal performance was robust. However, it was of interest to determine if the organic carbon profiling approach applied here could effectively discern operational changes and diagnose upsets. To examine whether this was the case, correlations were assessed between TOC, DOC, BDOC, NBDOC, and POC measurements and available water quality and operational data. The full results of this analysis are available in Appendix B: Subsection 2 and the supplemental spreadsheet, with a summary of key findings with respect to each stage of treatment summarized below.

Pilot Influent

Among non-upset sampling events, the typical percent DOC in the influent represented by BDOC ranged from 25.7% to 41.5% (SI Table A.1). However, during the methanol leaching events associated with SUMeth02 and SUMeth04, the pilot influent's DOC fraction was 55.9% and 56.2% BDOC, respectively. The conclusion that the percent influent BDOC was elevated by methanol was further supported by the fact that the NBDOC displayed minimal variability during these events and throughout operation (average NBDOC = 4.62 mg/L, SD = 0.459 mg/L, n=8). Overall, the BDOC analysis provided insights in upstream plant performance and upsets that more conventional TOC or DOC measurements would not be able to capture. Apart from the two sampling events influenced by methanol carry-over, influent concentrations of TOC, DOC, BDOC, and BDOC:NBDOC ratio were relatively uniform and did not correlate with any water quality parameters.

Floc-Sed

Floc-Sed operation was consistent throughout the testing period, except for one major upset event, SUTurbid09, where the aluminum chlorohydrate stock ran out and temporary adjustments were made to the coagulant dosing to compensate. The doses applied were found to be positively correlated with the change in POC concentration (Pearson: $r = 0.80$, $p<0.05$), with doses >32-35 mg/L resulting in increased POC in the filter effluent. Further, the change in the percent BDOC through Floc-Sed was negatively correlated with influent turbidity (Pearson: $r = -0.87$, $p<0.05$) and turbidity removal achieved through Floc-Sed (Pearson: $r = -0.92$, $p<0.05$). Correspondingly, the change in the NBDOC fraction was inversely correlated. Overall, the results demonstrate that elevated turbidity was effectively managed by the Floc-Sed stage, while BDOC fractions were disproportionately decreased when turbidity was elevated, relative to the NBDOC fraction.

Ozone

Change in the percent BDOC following ozone treatment was positively correlated with three operational indicators: ozone:TOC (Pearson: r =0.78, $p<0.05$), ozone:TOC with nitrite adjustment (Pearson: r =0.81, $p<0.05$), and applied ozone dose (Pearson: r =0.80, $p<0.05$). These strong correlations are consistent with the understanding that elevated ozone increases conversion of organic matter to a biodegradable form. Nitrite is a known ozone sink (1 mg/L of nitrite consumes 3.43 mg/L of ozone) and factoring in the corresponding reduction in overall oxidative capacity resulted in the strongest correlation.

Change in NBDOC concentration was found to be positively correlated with influent turbidity (Pearson: r =0.83, $p<0.05$), where reduction of NBDOC was less effective at higher influent turbidity. This suggests either that the influent POC is not very biodegradable, which is inconsistent with what was observed through Floc-Sed treatment, or, that POC imparted an ozone demand and diminished efficacy against NBDOC. The effluent NBDOC was also negatively correlated with the ozone:TOC dose (Pearson: r =-0.72, $p<0.05$), consistent with less effective conversion to BDOC at lower ozone dose.

BAC Filtration

BAC filtration relies on microbial treatment and thus was expected to be the most sensitive to changing conditions in the system. The BAC filtration units therefore were tested under a range of operating conditions (Table 2.1). At this point in the treatment train, flow was split between the BAC5 and BAC10 treatment units, enabling assessment of effects of EBCT. Both BAC units showed similar trends over time and indicated consistent nitrification, i.e., ammonia being removed and nitrite/nitrate being produced. In terms of organic carbon removal, BAC10 performed consistently better than BAC5. Also, correlation analysis of the BAC10 filter revealed a wider range of relationships between organic carbon profiles and water quality/operational data than the BAC5 filter. This suggests that the lower EBCT was not as stable, making it more vulnerable to high loading rates and shifting influent conditions, resulting in greater variability in the data.

Correlation analysis of the BAC10 unit revealed a strong and unexpected relationship with nitrogen species. Specifically, changes in BDOC concentrations were negatively correlated to changes in NH_4^+ concentration (Pearson: r= -0.94, $p<0.05$) and total inorganic nitrogen concentration (Pearson: r= -0.84, $p<0.05$) and positively correlated with effluent NO_3- concentration (Pearson: r =0.77, $p<0.05$) and its change during treatment (Pearson: r =0.8436, $p<0.05$). In other words, BDOC removal was enhanced when there was less removal of NH_4^+ and total inorganic nitrogen. The fact that the correlation was also strong with NO_3^- suggests that the driver of this relationship is not merely physical removal of NH_4^+ and total inorganic nitrogen, but biological nitrification. Percent removal of DOC was also correlated with the change (removal) in NH_4^+ concentration during BAC10 treatment (Pearson: r=0.77, $p<0.05$), temperature (Pearson: r=0.79, $p<0.05$), and monochloramine concentrations (Pearson: r=0.89, $p<0.05$). The latter relationship suggests that monochloramine, when coupled with dechlorination prior to BAC filtration, may have been more of an indicator of residual ammonia reducing competition for nitrogen between heterotrophs and nitrifiers, than acting as a disinfectant. However, when monochloramine was unquenched, BDOC removal was negatively impacted by the loading of monochloramine directly onto the filter (SSNoQuench-10).

In the BAC5 filter, the effluent BDOC concentration was also positively correlated with both monochloramine (Pearson: $r=0.87$, $p<0.05$) and ammonia concentrations (Pearson: $r=0.88$, $p<0.05$), while the ammonia concentration was further positively correlated with effluent TOC (Pearson: $r=0.85$, $p<0.05$). No operational parameter or conditions correlated strongly with treatment efficiency or bulk reductions of organic carbon.

Upset events provided the opportunity to assess the resilience of BAC filtration to changing operating conditions and to gain deeper insight into functionality of the units. During the higher BDOC loadings associated with the methanol carry-over event (SUMeth02 and SUMeth04), effluent TOC and DOC were elevated. However, net DOC removal was also elevated, indicating that the treatment train had the capacity to absorb the shock of the elevated BDOC load, with comparable BDOC across sampling events from the effluent of the BAC filters onward (SI Figure A.1). Effluent TOC and DOC were also elevated following Floc-Sed upset, which subjected the filters to increased influent turbidity (SUTurbid09). In this case, DOC removal was negatively impacted, but BAC10 was successfully able to continue treatment at a reduced rate, highlighting the robustness of BAC filters operated at longer EBCT. BAC5 was less successful at continued treatment, highlighting the susceptibility of shorter EBCT to operational variability and the need for more effective controls. There were also two instances of unintentional media loss from BAC5, following errors in backwashing. In both instances media was replaced and/or the influent flow rate was adjusted to normalize EBCT. After the first instance of media loss (04/01/2017), media from BAC10 was transferred into BAC5 prior to adjusting both filters flow rate to keep EBCT consistent. Following the transfer of media, SS-05 through SS-08 experienced increased removal of DOC and treatment efficiencies within BAC5 (Figure 2.2) by approximately 0.24 mg/L and 12%, respectively. BAC5's improved performance decreased following the upset of the Floc-Sed process (SUPolymerTurbid-09) and loading of monochloramine without quenching onto the filters (SSNoQunech-10). Conversely, BAC10's performance remained relatively unchanged.

Outside of process upsets, performance may also be linked to changes in operational parameters, as each sampling event with measurable changes in removal of the various organic carbon fractions was associated with changes in treatment, such as: chlorine loading onto the biofilters (SUMeth02 and SS10), high turbidity (SUTurbid09), and filter polymer addition (SUTurbid09 and SSPolymer10). Notably, filter polymer addition was associated with elevated percent POC in the BAC10 (variability in data likely masked this relationship for BAC5), potentially the result of increased biofilm development on polymer scaffolding. BDOC removal was most efficient during the stable operation period associated with SS05 through SS08, resulting in the lowest effluent TOC and DOC concentrations. Peak performance occurred during SS05 and SS06, when the temperature also peaked (>26 C°)(all other sampling events were < 18 F°), which likely relates to elevated microbial metabolism with elevated temperature. Just prior to SS07, temperature control was implemented at 15°C to eliminate this variable and create a 'worst' case for process operation (especially biodegradation kinetics), but BDOC removal remained high. Nutrients were also supplemented at this point, to ensure that the BAC filters were non nutrient limited, but it is not possible to evaluate the extent that each variable affected BDOC removal because they were changed at the same time.

Benchmarking of BDOC results to operational conditions and water quality data proved useful in explaining BAC performance variability. Other studies examining the fate of individual

trace organic contaminants via ozonation and biofiltration have similarly found that removal depends on a number of operational and environmental conditions, including: upstream ozone dose (Arnold et al., 2018; Gifford et al., 2018; Huber et al., 2003; Lee et al., 2013; Reungoat et al., 2012; Snyder et al., 2006; Sundaram & Pagilla, 2020; Ternes et al., 2002), upstream chemical additions (X. Liu et al., 2001; Urfer et al., 1997), temperature (Basu et al., 2016), media type (Arnold et al., 2018; Basu et al., 2016; LeChevallier et al., 1992), filter age (Sun et al., 2018), nutrient concentrations (Lauderdale et al., 2012) and EBCT (Arnold et al., 2018; Müller et al., 2017; Sundaram & Pagilla, 2020; Zearley & Summers, 2012). Our study found that BAC performance was strongly tied to EBCT and nitrification as well as the avoidance of operational upsets and limiting of microbial toxic compounds. Specifically, carbon removal during BAC5 was decreased during the Floc-Sed process upset (SUPolymerTurbid-09) and chlorine loading without quenching (SSNoQuench-10). The lower EBCT proved to be less resilient to non-ideal conditions. This was further supported by the result that organic carbon removal, the primary focus of BAC filtrations, was negatively correlated with ammonia oxidation and ultimately nitrification, potentially because of competition for ammonia. This highlights the sensitivity of biological treatment systems to factors such as temperature and nitrogen. In particular, NH_4+ availability appeared to be key, while monochloramine acted both as an inhibitory compound, but also as an indicator of the availability of NH_4+. Interestingly, BAC5's performance was improved and sustained following the transfer of media from BAC10 after media loss occurred prior to SS-05. This observation warrants additional research, as it is plausible that the transferred media seeded BAC5 with a more established, efficient microbial community resulting in improved DOC removal.

GAC Contacting

Organic carbon profiling corroborated the understanding that the GAC contactors are strongly influenced by performance of the BAC filters. Both GAC contactors exhibited elevated effluent TOC and DOC concentration during SUMeth02, SUMeth04, SUTurbid09, and SSPolymer10, highlighting the importance of optimizing upstream treatment for GAC contactors to perform optimally. Otherwise, both GAC contactors provided consistent organic carbon removal throughout the testing period.

BDOC analysis allowed estimation of NBDOC removal, i.e., the intended purpose of the GAC units. When combining the data from both GAC units, NBDOC removal negatively correlated with average GAC filter runtime (Pearson: $r=-0.66$, $p<0.05$) and temperature (Pearson: $r=-0.71$, $p<0.05$). GAC20's effluent NBDOC also negatively correlated with average filter runtime (Pearson: $r=-0.88$, $p<0.05$) and temperature (Pearson: $r=-0.87$, $p<0.05$), and additionally was positively correlated with NO_2^- concentration (Pearson: $r=0.88$, $p<0.05$). GAC contactor bed volumes were also negatively correlated with DOC removal (Pearson: $r=-0.62$, $p<0.05$) and treatment efficiency (Pearson: $r=-0.75$, $p<0.05$). These results are consistent with well-established principles of GAC exhaustion being driven by the total loading of organic carbon in the influent with time.

Similarly to BAC filters, BDOC removal by the GAC contactors was correlated with the fate of nitrogen species. Interestingly, change in percent NBDOC correlated with the change in NO_3^- and total inorganic nitrogen concentrations for both GAC units, when assessed independently and jointly (Pearson: $r=-0.89$, $p<0.05$and 0.81, $p<0.05$, for NO3- and total inorganic nitrogen, respectively; p and r values presented for the joint analysis, all values

reported in SI spreadsheets). Additionally, change in BDOC concentrations similarly correlated with changes in NO_3^- (Pearson: $r=0.92$, $p<0.05$) and total inorganic nitrogen concentration (Pearson: $r=0.87$, $p<0.05$). Treatment efficiency (percent DOC removal relative to influent DOC) also correlated with changes in NO_3^- (Pearson: $r=-0.93$, $p<0.05$) and total inorganic nitrogen (Pearson: $r=-0.79$, $p<0.05$) concentration. Relationships between BDOC and nitrogen removal suggest that, although the primary purpose of the GAC contactors is to remove NBDOC via sorption, there was some degree of biological removal occurring as well.

CONCLUSIONS

Here we refined and tested a BDOC protocol adapted for water reuse treatment trains employing biological removal via media with attached growth. Overall, this study demonstrated that BDOC testing can provide critical insight into drivers of organic carbon removal via carbon-based AWT, which can further inform efforts to improve reliability and performance for producing high quality water for potable reuse. Unlike traditional TOC and DOC analysis, BDOC analysis proved to be capable of yielding key insights into the fate of biodegradable and non-biodegradable fractions of dissolved organic carbon during various stages of AWT. BDOC analysis validated some fundamental expectations regarding carbon removal by various processes, but also generated new insights and hypotheses. For example:

- Modest BDOC removal was achieved during Floc-Sed, an effect previously reported for drinking water treatment (Volk et al., 2000), but BDOC was disproportionately reduced relative to NBDOC.
- Ozonation did not reduce DOC, but could reduce NBDOC (average reduction of 1.07 mg/L) by proportionally increasing BDOC, as expected, by 0.36-1.81 mg/L.
- BAC filtration effectively removed BDOC (average reduction of 0.82 mg/L), but did not significantly remove NBDOC (average reduction of 0.07 mg/L), as expected.
- GAC contacting was unexpectedly able to achieve 0.53 mg/L of average reduction in NBDOC, even with >20,000 treated bed volumes. However, inefficient BDOC removal during BAC treatment will detract from the ability of GAC to remove NBDOC.
- Improved performance was achieved at a longer EBCT. The longer EBCT treatment train (BAC10/ GAC20) outperformed the shorter EBCT treatment train in terms of overall effluent BDOC concentrations (0.76 mg/L versus 1.21 mg/L in the effluent, respectively).
- Benchmarking of BDOC measurements to operational conditions proved useful in identifying process upsets and explaining BAC performance variability. For example, BDOC analysis was effective at identifying elevated BDOC entering the pilot after methanol leaching upstream. Further, it provided insights into BAC filtrations resiliency to upsets by tracking the fate of these elevated concentrations and subsequent elevated removal rates during filtration. BAC5's performance was also improved following the inoculation of media from BAC10. Generally, BAC performance and BDOC removals were strongly tied to EBCT, the avoidance of elevated turbidity in Floc-Sed effluent, and limiting microbial toxic compounds (e.g. chlorine concentrations).
- Correlation of water quality data to BDOC measurements were also found to provide valuable insights into process performance. For example, significant relationships

were found between BDOC removal and nitrification/monochloramine concentrations in both BAC5 and BAC10. During ozonation, change in BDOC percentage was found to be positively correlated with multiple indicators for ozone dosing and ratios of ozone to TOC. With respect to NBDOC removal, GAC contactor performance was found to be negatively correlated with markers for contactor runtime while BDOC removal was correlated to nitrogen concentrations (similarly to the BAC filters), bringing to light unexpected biological removal of organic carbon occurring during GAC treatment.

Ideally, a carbon-based AWT train should produce finished water that is as biologically stable as possible, which can best be assessed by BDOC, rather than TOC or DOC. Also, BDOC analysis can be applied to ensure that upstream BDOC removal is maximized prior to GAC treatment, thus optimizing lifespan of the GAC and minimizing costly need for regeneration or replacement. It should be noted, however, that BDOC analysis requires substantially more time (15-30 days versus <1 hour) and effort (<1.5 hours versus <0.5 hours) than conventional TOC analysis. Therefore, its application should ideally be evaluated and completed in conjunction with conventional methods.

ACKNOWLEDGEMENTS

The authors would like to thank the operators and staff at Hampton Road Sanitation District's York River Treatment Plant and the SWIFT Research Center. Funding for this effort was provided in part by Water Research Foundation Unsolicited Award U1R16 (PI Pruden) and US Bureau of Reclamation grant R21 AC10162 (PI Pruden) 09/01/2020-08/31/2023 leveraged with additional financial and in-kind support from the Hampton Roads Sanitation District.

REFERENCES

1. Leverenz, H. L., Tchobanoglous, G. & Asano, T. Direct potable reuse: A future imperative. *J. Water Reuse Desalin.* **1**, 2–10 (2011).

2. Bernados, B. Reverse Osmosis for Direct Potable Reuse in California. *J. Am. Water Works Assoc.* **110**, 28–36 (2018).

3. Subramani, A. & Jacangelo, J. G. Treatment technologies for reverse osmosis concentrate volume minimization: A review. *Separation and Purification Technology* vol. 122 472–489 (2014).

4. Wenten, I. G. & Khoiruddin. Reverse osmosis applications: Prospect and challenges. *Desalination* **391**, 112–125 (2016).

5. Schimmoller, L. & Kealy, M. J. Fit for Purpose Water: The Cost of Overtreating Reclaimed Water. *WateReuse Res. Found. Proj. number 10-01* (2014).

6. Kim, S. D., Cho, J., Kim, I. S., Vanderford, B. J. & Snyder, S. A. Occurrence and removal of pharmaceuticals and endocrine disruptors in South Korean surface, drinking, and waste waters. *Water Res.* **41**, 1013–1021 (2007).

7. Ternes, T. A. *et al.* Removal of pharmaceuticals during drinking water treatment. *Environ. Sci. Technol.* **36**, 3855–3863 (2002).

8. Westerhoff, P., Yoon, Y., Snyder, S. & Wert, E. Fate of endocrine-disruptor, pharmaceutical, and personal care product chemicals during simulated drinking water treatment processes. *Environ. Sci. Technol.* **39**, 6649–6663 (2005).

9. Carlson, K. H. & Amy, G. L. BOM removal during biofiltration. *J. Am. Water Works Assoc.* **90**, 42–52 (1998).

10. Terry, L. G. & Summers, R. S. Biodegradable organic matter and rapid-rate biofilter performance: A review. *Water Research* vol. 128 234–245 (2018).

11. Yapsakli, K. & Çeçen, F. Effect of type of granular activated carbon on DOC biodegradation in biological activated carbon filters. *Process Biochem.* **45**, 355–362 (2010).

12. Kim, W. H., Nishijima, W., Shoto, E. & Okada, M. Pilot plant study on ozonation and biological activated carbon process for drinking water treatment. in *Water Science and Technology* vol. 35 21–28 (No longer published by Elsevier, 1997).

13. Hozalski, R. M., Goel, S. & Bouwer, E. J. TOC removal in biological filters. *J. Am. Water Works Assoc.* **87**, 40–54 (1995).

14. Vatankhah, H. *et al.* Evaluation of Enhanced Ozone-Biologically Active Filtration Treatment for the Removal of 1,4-Dioxane and Disinfection Byproduct Precursors from Wastewater Effluent. *Environ. Sci. Technol.* **53**, 2720–2730 (2019).

15. Nishijima, W. & Speitel, G. E. Fate of biodegradable dissolved organic carbon produced by ozonation on biological activated carbon. *Chemosphere* **56**, 113–119 (2004).

16. Rittmann, B. E. & Snoeyink, V. L. Achieving biologically stable drinking water. *J. / Am. Water Work. Assoc.* **76**, 106–114 (1984).

17. Garner, E., Zhu, N., Strom, L., Edwards, M. & Pruden, A. A human exposome framework for guiding risk management and holistic assessment of recycled water quality. *Environmental Science: Water Research and Technology* vol. 2 580–598 (2016).

18. Basu, O. D., Dhawan, S. & Black, K. Applications of biofiltration in drinking water treatment - a review. *J. Chem. Technol. Biotechnol.* **91**, 585–595 (2016).

19. Hallé, C., Huck, P. M. & Peldszus, S. Emerging Contaminant Removal by Biofiltration: Temperature, Concentration, and EBCT Impacts. *J. Am. Water Works Assoc.* **107**, E364–E379 (2015).

20. Gerrity, D. *et al.* Microbial community characterization of ozone-biofiltration systems in drinking water and potable reuse applications. *Water Res.* (2018) doi:10.1016/j.watres.2018.02.023.

21. Peterson, E. S. & Summers, R. S. Removal of effluent organic matter with biofiltration for potable reuse: A review and meta-analysis. *Water Res.* **199**, 117180 (2021).

22. Üstün, G. E., Solmaz, S. K. A., Çiner, F. & Bažkaya, H. S. Tertiary treatment of a

secondary effluent by the coupling of coagulation–flocculation–disinfection for irrigation reuse. *Desalination* **277**, 207–212 (2011).

23. Volk, C. *et al.* Impact of enhanced and optimized coagulation on removal of organic matter and its biodegradable fraction in drinking water. *Water Res.* **34**, 3247–3257 (2000).

24. Volk, C. J. & Lechevallier, M. W. Effects of Conventional Treatment on AOC and BDOC Levels. *J. Am. Water Works Assoc.* **94**, 112–123 (2002).

25. Vaidya, R., Wilson, C. A., Salazar-Benites, G., Pruden, A. & Bott, C. Implementing Ozone-BAC-GAC in potable reuse for removal of emerging contaminants. *AWWA Water Sci.* **2**, e1203 (2020).

26. Khan, E., Babcock, R. W., Viriyavejakul, S., Suffet, I. H. & Stenstrom, M. K. Biodegradable dissolved organic carbon for indicating wastewater reclamation plant performance and treated wastewater quality. *Water Environ. Res.* **70**, 1033–1040 (1998).

27. van der Kooij, D., Visser, A. & Hijnen, W. A. M. Determining the concentration of easily assimilable organic carbon in drinking water. *J. Am. Water Works Assoc.* **74**, 540–545 (1982).

28. Kooij, D. van der. Assimilable Organic Carbon as an Indicator of Bacterial Regrowth. *J. Am. Water Works Assoc.* **84**, 57–65 (1992).

29. Servais, P., Anzil, A. & Ventresque, C. Simple Method for Determination of Biodegradable Dissolved Organic Carbon in Water. *Appl. Environ. Microbiol.* **55**, (1989).

30. Escobar, I. C. & Randall, A. A. Assimilable organic carbon (AOC) and biodegradable dissolved organic carbon (BDOC): Complementary measurements. *Water Res.* **35**, 4444–4454 (2001).

31. LeChevallier, M. W., Shaw, N. E., Kaplan, L. A. & Bott, T. L. Development of a Rapid Assimilable Organic Carbon Method for Water. *Appl. Environ. Microbiol.* **59**, (1993).

32. Volk, C. J. & LeChevallier, M. W. Assessing biodegradable organic matter. *J. Am. Water Works Assoc.* **92**, 64–76 (2000).

33. Page, D. & Dillon, P. Measurement of the biodegradable fraction of dissolved organic matter relevant to water reclamation via aquifers. in *National Research Flagships* 33 (2007).

34. Park, S.-K., Pak, K.-R., Choi, S.-C. & Kim, Y.-K. Evaluation of Bioassays for Analyzing Biodegradable Dissolved Organic Carbon in Drinking Water. *J. Environ. Sci. Heal. Part A* **39**, 103–112 (2004).

35. Trulleyová, Š. & Rulík, M. Determination of biodegradable dissolved organic carbon in waters: Comparison of batch methods. *Sci. Total Environ.* **332**, 253–260 (2004).

36. Vaidya, R. *et al.* Pilot Plant Performance Comparing Carbon-Based and Membrane-Based Potable Reuse Schemes. *Environ. Eng. Sci.* **36**, 1369–1378 (2019).

37. Sun, Y. *et al.* Mathematical modeling of biologically active filtration (BAF) for potable water production applications. *Water Res.* **167**, 115128 (2019).

38. Vaidya, R., Wilson, C. A., Salazar-Benites, G., Pruden, A. & Bott, C. Factors affecting removal of NDMA in an ozone-biofiltration process for water reuse. *Chemosphere* **264**, 128333 (2021).

39. Pruden, A., Bott, C., Blair, M. F., Miller, J. H. & Vaidya, R. Characterization of Organic Carbon and Microbial Communities for the Optimization of Biologically-Active Carbon (BAC) Filtration for Potable Reuse. *Water Res. Found. Proj. 4872/U1R16 Final Report.* 305-undefined (2020).

40. Team, Rs. Citing RStudio – RStudio Support. *RStudio: Integrated Development for R. RStudio* https://support.rstudio.com/hc/en-us/articles/206212048-Citing-RStudio (2020).

41. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).

42. Zhiteneva, V., Ziemendorf, É., Sperlich, A., Drewes, J. E. & Hübner, U. Differentiating between adsorption and biodegradation mechanisms while removing trace organic chemicals (TOrCs) in biological activated carbon (BAC) filters. *Sci. Total Environ.* **743**, 140567 (2020).

43. Wu, H. & Xie, Y. F. Effects of EBCT and Water Temperature on HAA Removal using BAC. *J. Am. Water Works Assoc.* **97**, 94–101 (2005).

44. de Vera, G. A. & Wert, E. C. Using discrete and online ATP measurements to evaluate regrowth potential following ozonation and (non)biological drinking water treatment. *Water Res.* **154**, 377–386 (2019).

45. Thayanukul, P., Kurisu, F., Kasuga, I. & Furumai, H. Evaluation of microbial regrowth potential by assimilable organic carbon in various reclaimed water and distribution systems. *Water Res.* **47**, 225–232 (2013).

46. Joret, J. C., Levi, Y. & Volk, C. Biodegradable dissolved organic carbon (BDOC) content of drinking water and potential regrowth of bacteria. in *Water Science and Technology* vol. 24 95–101 (IWA Publishing, 1991).

47. Hijnen, W. A. M. *et al.* Slowly biodegradable organic compounds impact the biostability of non-chlorinated drinking water produced from surface water. *Water Res.* **129**, 240–251 (2018).

48. Jarusutthirak, C. & Amy, G. Understanding soluble microbial products (SMP) as a component of effluent organic matter (EfOM). *Water Res.* **41**, 2787–2793 (2007).

49. Liu, J. lin, Li, X. yan, Xie, Y. feng & Tang, H. Characterization of soluble microbial products as precursors of disinfection byproducts in drinking water supply. *Sci. Total Environ.* **472**, 818–824 (2014).

50. Carlson, K. H. & Amy, G. L. The importance of soluble microbial products (SMPs) in

biological drinking water treatment. *Water Res.* **34**, 1386–1396 (2000).

51. Hong, S., Xian-chun, T., Nan-xiang, W. & Hong-bin, C. Leakage of soluble microbial products from biological activated carbon filtration in drinking water treatment plants and its influence on health risks. *Chemosphere* **202**, 626–636 (2018).

52. Bal Krishna, K. C., Sathasivan, A. & Chandra Sarker, D. Evidence of soluble microbial products accelerating chloramine decay in nitrifying bulk water samples. *Water Res.* **46**, 3977–3988 (2012).

53. Laspidou, C. S. & Rittmann, B. E. A unified theory for extracellular polymeric substances, soluble microbial products, and active and inert biomass. *Water Research* vol. 36 2711–2720 (2002).

54. Namkung, E. & Rittmann, B. E. Soluble microbial products (SMP) formation kinetics by biofilms. *Water Res.* **20**, 795–806 (1986).

55. Alvarado, A., West, S., Abbt-Braun, G. & Horn, H. Hydrolysis of particulate organic matter from municipal wastewater under aerobic treatment. *Chemosphere* **263**, 128329 (2021).

56. Levine, A. D., Tchobanoglous, G. & Asano, T. Size distributions of particulate contaminants in wastewater and their impact on treatability. *Water Res.* **25**, 911–922 (1991).

57. Rickert, D. A. & Hunter, J. V. General nature of soluble and particulate organics in sewage and secondary effluent. *Water Res.* **5**, 421–436 (1971).

58. Steirer, M. A. *et al. City of San Diego Water Purification Demonstration Project Project Report*. https://www.sandiego.gov/sites/default/files/legacy/water/purewater/pdf/projectreports/wpdpfinalprojectreport.pdf (2013).

59. Servais, P., Billen, G. & Bouillot, P. Biological Colonization of Granular Activated Carbon Filters in Drinking-Water Treatment. *J. Environ. Eng.* **120**, 888–899 (1994).

CHAPTER 3: SUCCESSION OF THE RESISTOME THROUGH AN ADVANCED OXIDATION PROCESS-BIOLOGICALLY ACTIVE CARBON FILTRATION-BASED TREATMENT TRAIN FOR POTABLE REUSE

Matthew F. Blair, Peter Vikesland, Amy Pruden, and Charles Bott

ABSTRACT

The ability for water reuse applications to alleviate stress on conventional water resources is being realized through the increased implementation of advanced water treatment trains to produce water suitable for potable reuse. Further, treatment trains that rely on biological removal of contaminants have gained additional popularity because they are capable of breaking down pollutants into less harmful chemicals, are less costly, and overall more sustainable when compared to membrane-based treatment systems. However, with increased biological activity comes a need to better understand the fate of microbial contaminants of emerging concern (CECs) within these systems. Of particular concern is antibiotic resistance, due to the nature of biological treatment, stress from upstream oxidative processes, pathogens associated with wastewater, and increased contact with resilient chemical CECs (e.g. antibiotics). Here, we utilized shotgun metagenomic sequencing to characterize the fate of antibiotic resistance genes (ARGs) carried across the microbial community (i.e., the resistome) throughout a carbon-based potable reuse train. It was found that the resistome was heavily shaped by individual treatment processes and the corresponding microbial communities. Conclusions about the fate of ARGs were dependent on whether evaluated in terms of relative or calculated absolute abundances. Notably, disinfection processes provided effective reduction of calculated absolute abundances at the expense of elevated relative abundances. Biologically-active carbon filtration also proved to be an important intermediate process, as effluents were typically enriched in both relative and absolute abundances of ARGs. Overall, the treatment train was found to effectively reduce the loading of total ARGs with comparatively few clinically-relevant ARGs found throughout the whole treatment train. The findings of this study helped to ensure that water reuse treatment trains can adequately address concerns about antibiotic resistance.

INTRODUCTION

Challenging the conventional paradigm of wastewater treatment with a framework focused on water reuse requires that a greater emphasis be placed on understanding the impact of reuse treatment trains on emerging public health concerns. In particular, design of water reuse treatment trains needs to consider the fate of biological contaminants of emerging concern, antibiotic resistance genes (ARGs) chiefly among them, in addition to the obvious need to demonstrate effective, robust removal of viral and bacterial pathogens[1]. Establishing an understanding of microbial dynamics will be especially important for processes that harness biological treatment, such as biologically active filtration (BAF). BAF is gaining popularity in large part because it does not require membrane filtration, which is costly, energy-intensive, and produces an undesirable brine waste stream. Instead, BAF-based treatment trains rely on the implementation of multiple barriers for biological contaminant removal, including microbially destructive AOP (e.g. ozonation, and UV-H2O2), enhanced solid liquid separation (e.g. coagulation, flocculation, and sedimentation), and disinfection. Multiple treatment barriers are especially important in direct potable reuse applications, which do not incorporate natural

barriers inherent in receiving environment and thus may short circuit exposure routes. In such scenarios, it is critical to expand biological monitoring beyond traditional bacterial indicators for more robust, comprehensive monitoring with higher standards[2].

Antibiotic resistance is of particular concern due to its negative impact on public health and the fact that wastewater treatment plants are recipients of bacteria (both pathogenic and nonpathogenic), antibiotic resistant bacteria, residual antibiotics, metals and various other compounds that potentially play a role in the dynamics of antimicrobial selection, proliferation, and dissemination. Further, wastewater, drinking water, and advanced treatment processes can provide a valuable critical control point in preventing the spread of environmental antimicrobial resistance (AMR)[3]. Thus, numerous studies have been conducted to explore the fate of specific ARGs through conventional wastewater treatment processes[4–7] as well as characterize the resistome (i.e. total ARGs carried across the microbial community)[8] associated with activated sludge processes[9,10], raw sewage[11], and wastewater treatment plant effluents[12–14]. Fewer studies have looked at the specific impact of wastewater treatment processes on resistomes[15,16] or with respect to various types of water reuse[17–19], leaving substantial knowledge gaps and a need for comprehensive analysis of the dynamics related to antibiotic resistance and various mediating factors (antibiotics, metals, mobile genetic elements, etc.) within potable reuse treatment trains. Those employing biologically driven processes are of specific interest because of the potential for ARG selection and horizontal transfer[20] within the complex bacterial communities responsible for treatment efficacy.

Though beneficial, more conventional microbial methods, such as culture- and quantitative polymerase chain reaction (qPCR), fail to comprehensively characterize the extensive, complex interplay associated with the resistome and various treatment technologies. Ideally, AMR characterization should provide as much breadth and depth as possible, yielding comprehensive understanding of various dimensions by which ARGs are influenced by various treatment processes[21]. Shotgun metagenomic sequencing holds promise as a robust means to provide this comprehensive characterization by directly accessing a vast assortment of ARGs, shedding light on implications for AMR and the factors that contribute to it within water reuse applications, while also informing development of potential mitigation strategies, policy, and best practices. However, metagenomic sequencing is not without its drawbacks, as cost and complexity of analysis have limited widespread application, especially in longitudinal studies. Concerns also exist about metagenomics' quantitative capacity[22]. Ideally, metagenomic sequencing can be used to better understand the dynamics of AMR within advanced treatment trains.

Here we utilized shotgun metagenomic sequencing to characterize the behavior of ARGs throughout a carbon-based potable reuse train employing coagulation, flocculation, sedimentation, ozonation, BAC filtration, GAC contacting, and UV disinfection. Spatial and temporal shifts in the resistome were related to fluctuations in process performance and changes in operational conditions. Specifically, the objectives of this work were to: 1) compare the effects of physical, chemical, and biological treatment processes on the establishment of the resistome and determine the extent that each component of the treatment train provides a barrier to ARG dissemination, 2) characterize the for BAC and GAC treatment following AOP treatment to enrich bacteria in the community carrying ARGs, 3) determine if disinfection (i.e., ozone and UV disinfection) effectively reduces ARGs, 4) assess differences between total ARGs and those

that are of particularly high clinical relevance, and 5) classify core and discriminatory ARGs unique to each treatment process. The findings of this study advance fundamental understanding of the fate of ARGs and resistome dynamics occurring within carbon-based advanced water treatment trains intended for potable reuse.

METHODS

Site Description, Sample Collection, and Preservation

Samples were collected from each stage of the 4.3 gpm SWIFT Pilot Scale facility at the York River Wastewater Treatment Plant (Figure 3.1), over approximately an 18-month period. The SWIFT Pilot treats tertiary (denitrified secondary) effluent and employs: coagulation-flocculation-sedimentation (Floc-Sed), ozonation, BAC filtration, GAC contacting, and ultraviolet (UV) disinfection. To assess the impact of EBCT, the pilot flow was split between a relatively low EBCT stream of 5 min for the BAC stage (BAC5) and 10 min for the GAC stage (GAC10) or a high EBCT stream of 10 min for the BAC stage (BAC10) and 20 min for the GAC stage (GAC20). In addition to the treatment processes themselves, various chemicals and operational conditions were also adjusted periodically with major changes and upset conditions outlined in SI Table A.1. Notably, the treatment processes were subjected to changes in targeted ozone dosing, suppressed temperatures, changes in chlorine quenching prior to BAF, the addition of supplemental nutrients and filter aid polymer prior to BAF treatment, and a few upset conditions (e.g., media loss on the BAC filters, and elevated turbidity following coagulation, flocculation, and sedimentation). Additional details on SWIFT pilot operation during this study period are available in prior publications[23–27] and the SI.

Triplicate samples were collected on site and transported on ice to Virginia Tech in sterile 1-L polypropylene bottles. Processing of collected samples occurred within 24 h of collection and included: filter concentration onto 0.22- µm mixed cellulose esters membrane filter (Millipore, Billerica, MA), fragmentation via sterilized tweezers, and storage at – 20°C. The exact sample volume that was filtered varied with total suspended solids content of the sample as the solids clog the filter. Exact sample volumes were recorded at the time of filtering, but typically ranged from 0.2 (Pilot influent) to 1 liter (all other sampling locations). DNA was extracted from filters using a FastDNA SPIN Kit for Soil (MP Biomedicals, Solon, OH) according to the manufacturer's instructions. Extracts were then stored at -20 °C (short term), or -80 °C (long term) prior to downstream molecular analysis.

Shotgun Metagenomic Sequencing and Analysis

72 samples, collected over 10 sampling events, were selected from the entire sampling campaign, and submitted for shotgun metagenomic sequencing over three individual lanes. DNA concentrations were quantified using the Qubit dsDNA HS-assay (Invitrogen™ Qubit™ 3 Fluorometer, Waltham, MA) and samples with concentrations below 0.2 ug/L were not considered for metagenomic sequencing. Before being submitted for sequencing, DNA extracts were pooled across replicates (each sample at equal mass) to achieve the target concentration. Sequencing was completed by The Scripps Research Institute (The Scripps Research Institute, San Diego, CA, United States) on an Illumina NextSeq500 (Illumina, San Diego, CA, United States) using 2x75 paired end reads and the NEB Ultra II library prep, which is suitable for

samples with DNA masses down to 1ng. This library was selected because of the low concentration of DNA typical of the ozone and UV effluent samples.

Reads were uploaded, read matched, and annotated using the Metastorm platform [28] and default parameters. The target depth corresponded to approximately 17 million reads per samples with additional metagenomic read statistics detailed in SI Figure C.1 through C.4. Current analysis utilized a manually curated version of CARD v2.0.1[29] for annotation of bulk ARGs into resistance classes. Genes were normalized using the Silva rRNA database[30] for 16S rRNA genes (i.e., ARG copies per copies of 16S rRNA genes generated from metagenomic data)[31] to calculate relative abundances of ARGs. Calculated absolute abundances (i.e., ARG copies/mL) were calculated by multiplying each sample's 16S rRNA normalized ARG copies by their respective 16S rRNA gene abundance, quantified by qPCR targeting the 16S rRNA gene[18]. Further annotation was also provided for ARGs identified as being 'clinically relevant' if they conferred resistance to antibiotics used to treat infections in humans. The manual curation utilized in this study and the list of clinically relevant genes are available in SI spreadsheet.

Data and Statistical Analysis

Bray-Curtis dissimilarities were generated in R[33] to compare the resistome between stages of treatment, sampling events, and changes in operational conditions. Associated statistical testing utilized the analysis of similarities (ANOSIM) for categorical variables and ADONIS test for continuous variables, both from the vegan package[34]. R was also used to assess differences in abundances between stages of treatment using a paired Wilcoxon rank-sum test and differences between sampling events were tested using Dunn's T-test. Results were also reported for antibiotic resistance classes that experienced consistent reductions or enrichments, which was defined as occurring in greater than 70% of comparisons. Statistical significance was set at $p < 0.05$, with full statistical reporting presented in SI spreadsheet.

Procrustes analysis was completed using the vegan R package to compare the resistome, calculated using shotgun metagenomics, and the microbial community structure determined via 16S rRNA gene amplicon sequencing. Amplicon sequencing was carried out through PCR amplification of DNA extracts using the 515f/926r universal bacterial/archaeal primer set targeting the V4-V5 region of the 16S rRNA gene. The Fralin Life Sciences Institute Genomic Sequencing Center at Virginia Tech (Blacksburg, VA) sequenced the amplicons on an Illumina MiSeq using a 250 cycle paired end protocol while sequenced reads were processed using the QIIME2 pipeline and annotated using the Greengenes database (May 2013 release).

The core resistome was defined to include any ARG with a non-zero relative abundance that was found in all sampling events for any given treatment process. Discriminatory ARGs were identified by LEfSe (Linear discriminant analysis Effect Size) [35] analysis for both raw gene counts and 16S rRNA normalized abundances (presented). All figures were generated using ggplot2 [36] in R.

RESULTS

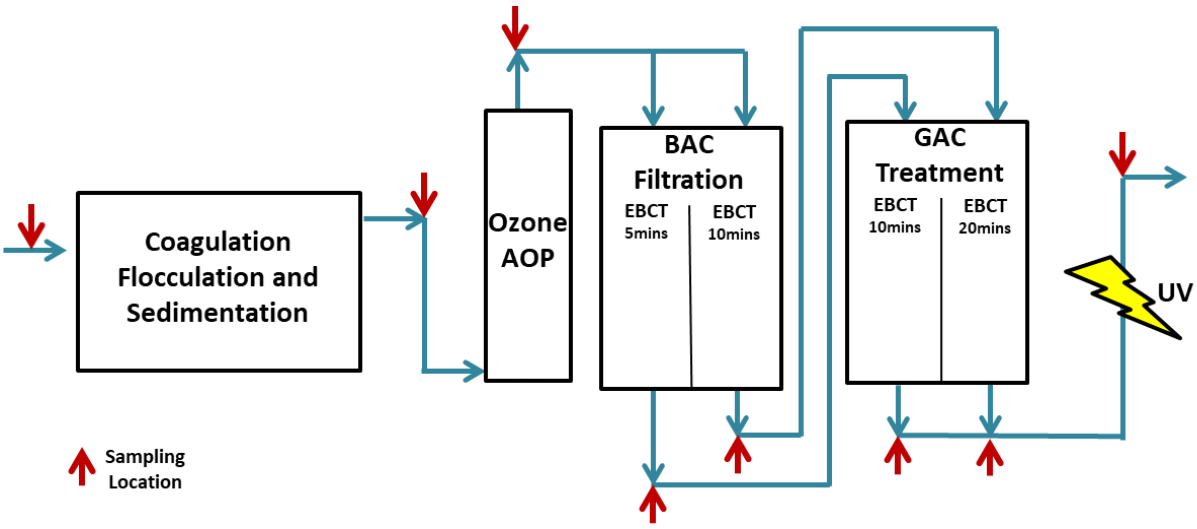


Figure 3.1: HRSD SWIFT Pilot Facility's process schematic. Red arrows indicate sampling locations. The pilot influent consisted of effluent from the York River wastewater treatment plant at a flowrate of 4.3 gallons per minute.

Trends in ARG Abundance Detected throughout the Treatment Train

Figure 3.2a reports relative abundances (ARG copies/16S rRNA gene copies) for total annotated ARGs in the CARD database and those sub-classified as clinically-relevant grouped by antibiotic resistance class and averaged by treatment process. Figure 3.2b reports corresponding calculated absolute abundances [Log(ARG gene copies per mL)]. Additional figures wherein ARGs are not averaged by treatment process can be found in Appendix D: Subsection 1. From Figure 3.2, on average the relative abundance of ARGs remained relatively unchanged during the floc-sed process. A slight increase (8.1%) was subsequently observed following ozonation. At the same time, both floc-sed and ozonation resulted in bulk reductions in bacteria, as reflected by total 16S rRNA genes (quantitated via qPCR). As a result, the calculated absolute abundances of all quantified ARGs also decreased. The average relative abundance of ARGs increased within all BAC and GAC units relative to ozone effluent, with values being slightly higher in the BAC effluents. When accounting for changes in 16S rRNA genes through the measure of average calculated absolute abundances, similar trends were seen in BAC and GAC effluents relative to ozone where ARG were found in slightly higher concentrations throughout. However, the overall reduction of bacterial markers compared to influent samples mitigated the impact of elevated relative abundances and an overall decrease in calculated abundances of ARGs in the BAC and GAC samples compared to the influent. Further, higher EBCTs resulted in elevated calculated absolute abundances of ARG during both BAC and GAC treatments. Finally, UV disinfection experienced similar trends in measures of ARG abundances as ozonation, where relative abundances of ARGs were increased while net reduction in bacterial markers significantly decreased the overall concentration of ARGs. Clinically-relevant ARGs were also evaluated and it was found that their relative abundances comprised an extremely small (average of 2.1%) portion of total identified ARGs.

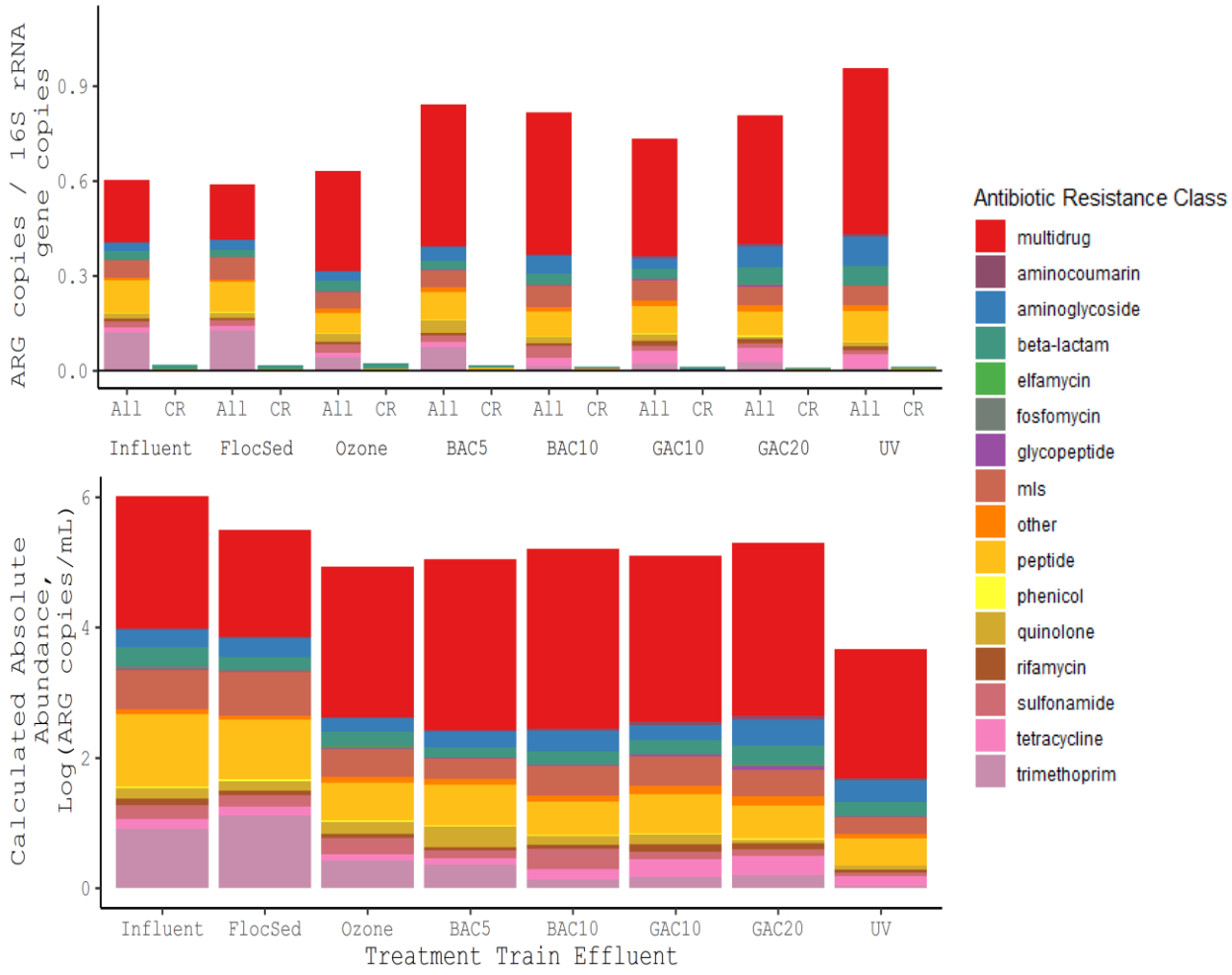


Figure 3.2: (A) Average relative abundance of total annotated ARG (i.e., average of all detected ARGs) relative abundances normalized to 16S rRNA gene copies and grouped by antibiotic resistance class at each treatment process. Clinically relevant relative abundances are similarly provided. (B) Average calculated absolute abundances [Log(ARG copies/mL)] for bulk classified ARGs; relative abundances multiplied by 16S rRNA gene copies quantified via qPCR. MLS corresponds to resistance to macrolide, lincosamides, and streptogramins while 'multidrug' represents resistances to at least two resistance classes. Resistance class shading is scaled proportionally to the total abundance of each stacked bar to ease comparison.

In general, the majority of annotated ARGs identified were found to confer resistance to multidrug, mls, and peptide classes of antibiotics. ARGs associated with aminocoumarin, aminoglycoside, beta-lactam, elfamycin, fosfomycin, glycopeptide, phenicol, quinolone, rifamycin, sulfonamide, tetracycline, and trimethoprim resistance classes were also identified. Clinically-relevant ARGs were found to be associated with beta-lactam, glycopeptide, peptide, and quinolone resistance classes. More specific changes in ARG class abundances are discussed in later sections.

Resistome Analysis, Non-metric Multidimensional Scaling (NMDS)

The resistome for each sampled treatment process was assessed via Bray Curtis dissimilarity metrics to identify broad differences in ARG profiles between treatment stages and overall variability within treatments. Selected NDMS plots graphically present these relationships for both measures of calculated absolute abundance (Figure 3.3) and relative abundance (SI Figure D.1). From Figure 3.3a, the resistomes clustered by their associated treatment trains (ANOSIM, p-value<0.05, r-stat = 0.67|0.42) more so than any other tested factor for both calculated absolute abundances and relative abundances. Similar statistically significant results were seen for comparisons between influent:ozone (ANOSIM, p-value<0.05, r-stat = 0.58|0.22), influent:UV (ANOSIM, p-value<0.05, r-stat = 0.98|0.64), BAC5:GAC10 (ANOSIM, p-value<0.05, r-stat = 0.16|0.20), and GAC:UV (ANOSIM, p-value<0.05, r-stat = 0.88|0.27) for both calculated absolute abundances and relative abundances, respectively. Similar relationships were seen between BAC filters and GAC contactors (Figure 3.3b, ANOSIM, p-value<0.05, r-stat = 0.16|0.14) and between the pilot influent and both disinfection processes (Figure 3.3c, ANOSIM, p-value<0.05, r-stat = 0.73|0.40). Interestingly, regardless of distinctions between calculated absolute abundances and relative abundances, no significant differences were found between the resistomes associated with influent:flocsed, BAC5:BAC10, and GAC10:GAC20. Associations were also found between flocsed:ozone (ANOSIM, p-value<0.05, r-stat = 0.36), ozone:BAC5 (ANOSIM, p-value<0.05, r-stat = 0.21), ozone:BAC10 (ANOSIM, p-value<0.05, r-stat = 0.45) and BAC10:GAC20 (ANOSIM, p-value<0.05, r-stat = 0.15) only when considering calculated absolute abundances.

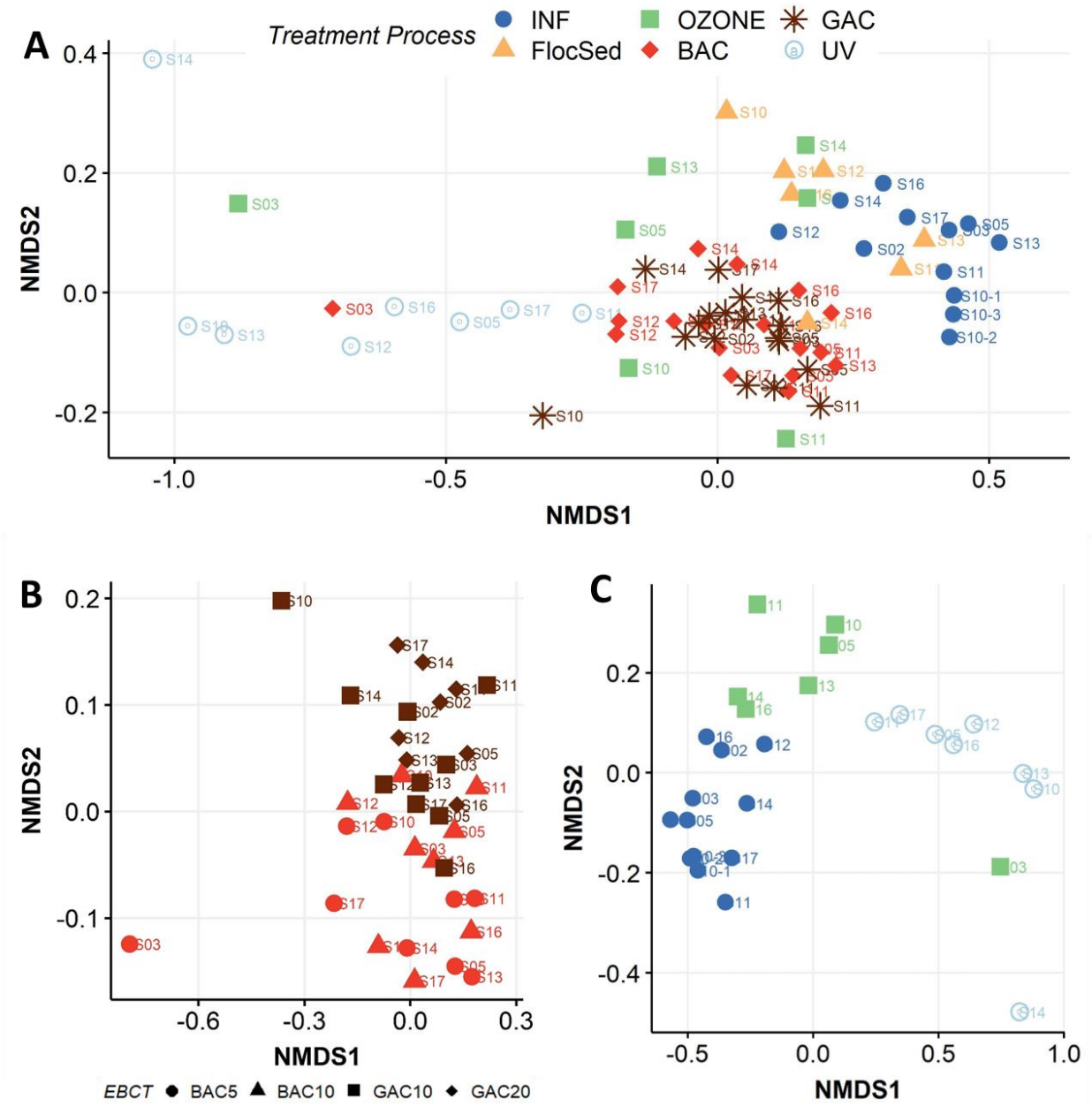


Figure 3.3: NMDS analysis of resistome (based on calculated absolute abundances) for (A) all treatment processes (ANOSIM, p-value<0.05, r-stat = 0.67), (B) BAC filtration and GAC contracting at tested EBCTs (ANOSIM, p-value<0.05, r-stat = 0.16), and (C) disinfection processes (ozone effluent and UV effluent) referenced to treatment train influent (ANOSIM, p-value<0.05, r-stat = 0.73). Similar figures for 16S rRNA normalized relative abundances are presented in the SI.

Comparison of the Resistome to Microbial Community Structure

Multiple Procrustes analysis were performed on the resistome and taxonomic NMDSs of all samples from each stage of treatment. These tests indicated statistically significant relationships between the resistome and microbial community structure when compared via

presence/absence (p-value < 0.05, R = 0.782), calculated absolute abundance (p-value < 0.05, R = 0.549), and counts (p-value < 0.05, R = 0.508) based analyses. For comparison to the NDMS plots, the calculated absolute abundance-based analysis is graphically presented in Figure 3.4, with residuals graphically presented in SI Figure D.4.

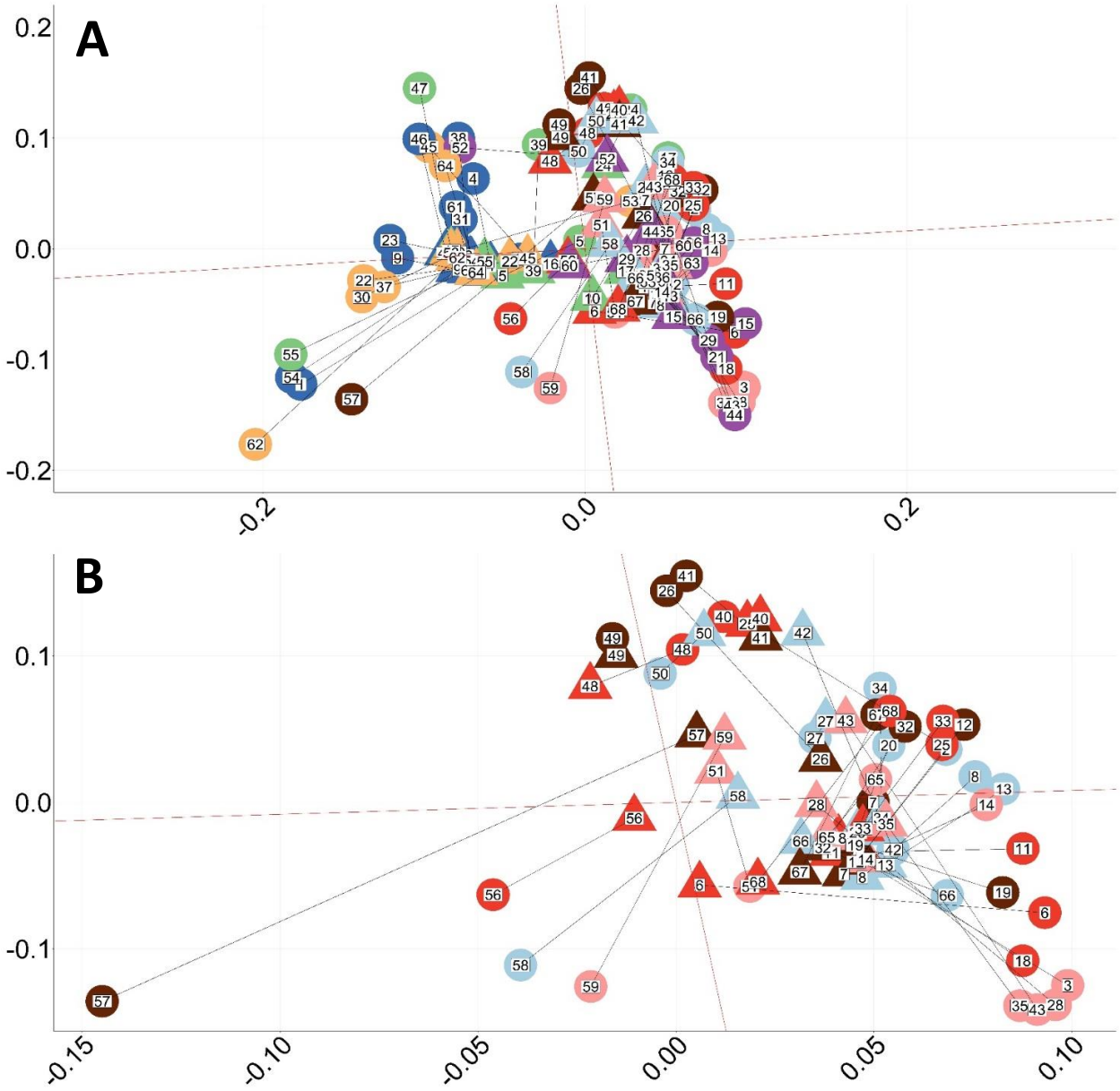


Figure 3.4: Procrustes analyses were performed on the resistome and taxonomic NMDSs of all samples from each stage of treatment. Tests resulted in statistically significant relationships occurring between the resistome and microbial community structure when compared via presence/absence (p-value < 0.05, R = 0.782), calculated absolute abundance (p-value < 0.05, R = 0.549) which is displayed above to be consistent with presented NDMS plots. Residuals associated with the Procrustes analysis are graphically presented in SI Figure D.4.

Assessment of Classified ARG Enrichment and Removal throughout Treatment Processes

Figure 3.5 provides a graphical representation of each antibiotic resistance classes' magnitude of change for both measures of relative and calculated absolute abundances, throughout each individual treatment process and the full treatment train. Summary statistics related to Figure 3.5 can be found in the SI spreadsheet with complete reporting of statistically significant comparisons.

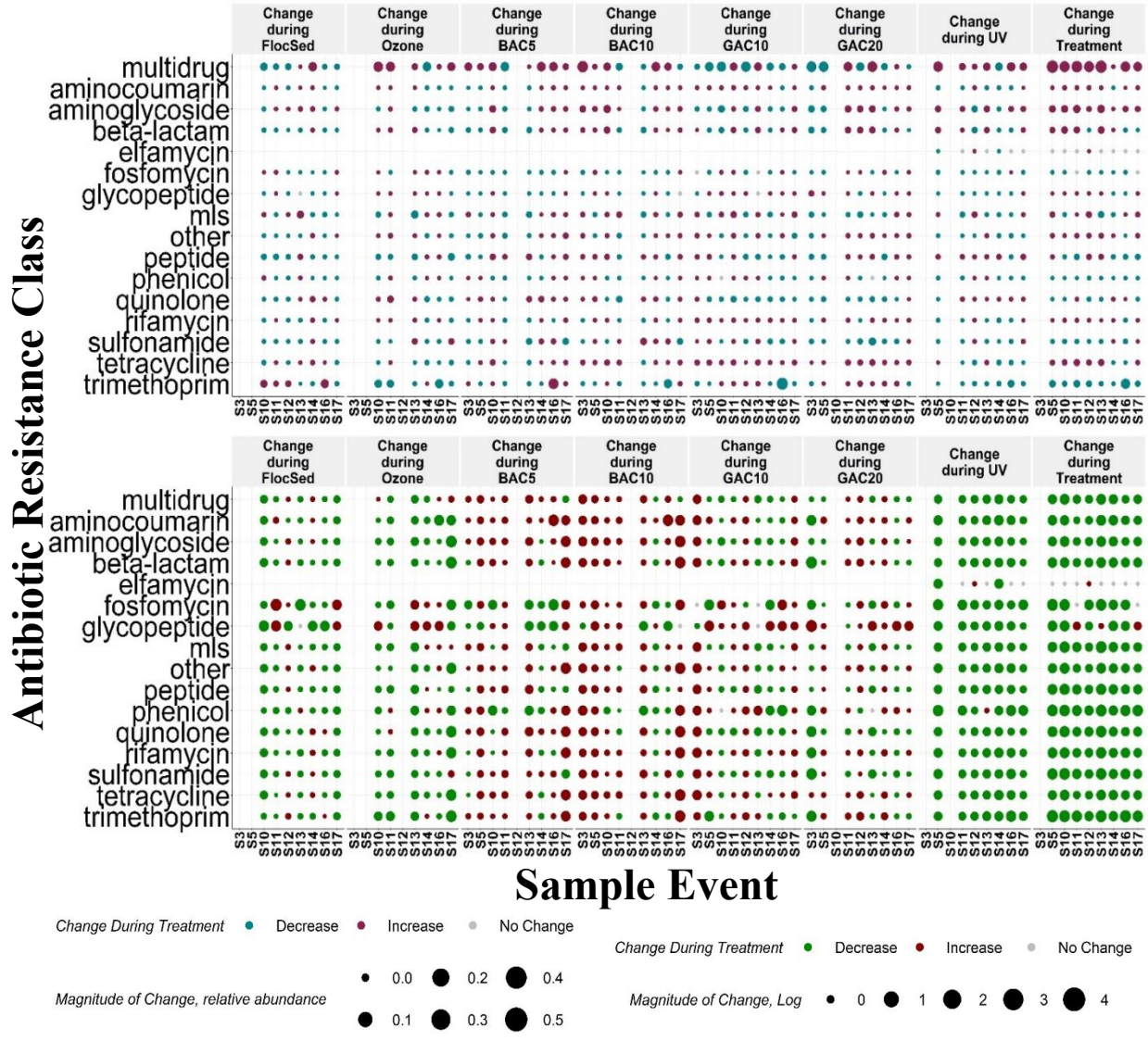


Figure 3.5: Magnitude of change for each antibiotic resistance class at each sample location and throughout the entire treatment train for (A) difference in 16S rRNA normalized abundances and (B) log change of the calculated absolute abundances.

When considering the impact of the flocsed process on the relative abundance of total annotated ARGs, on average, antibiotic resistance classes were reduced in 69.1% of cases.

Assessment of each individual antibiotic resistance class indicated that sulfonamide, phenicol, peptide, other, beta-lactam, and multidrug all reduced in relative abundance, in >70% of cases, while only the aminoglycoside class of ARGs was enriched. However, these differences were not statistically significant (Wilcoxon test). Due to flocsed's ability to successfully reduce the 16S rRNA gene marker for total bacterial number, assessment of calculated absolute abundances identified an elevated average reduction where 74.3% of all antibiotic resistance classes were removed. Of all 15 identified antibiotic classes, only fosfomycin, and glycopeptide were not reduced in greater than 70% of sampling events. Of the 13 antibiotic resistance classes that were consistently reduced, beta-lactam, multi-drug, peptide, quinolone, and rifamycin were found to be significantly different (Wilcoxon, p-value<0.5).

The ozone process was found to provide average reductions in 51.1% and 75.6% of sampling events for antibiotic resistance classes' relative and calculated absolute abundances, respectively. The former, indicated glycopeptide (Wilcoxon, p-value<0.5), and multidrug resistance classes to be enriched in greater than 70% of sampling events while the aminoglycoside, tetracycline, and trimethoprim (Wilcoxon, p-value<0.5) classes experienced consistent relative abundance reductions. When considering the changes in calculated absolute abundances, no resistance classes increased as a result of ozonation, while aminocoumarin (Wilcoxon, p-value<0.5), aminoglycoside (Wilcoxon, p-value<0.5), beta-lactam (Wilcoxon, p-value<0.5), other (Wilcoxon, p-value<0.5), peptide (Wilcoxon, p-value<0.5), quinolone, rifamycin (Wilcoxon, p-value<0.5), sulfonamide (Wilcoxon, p-value<0.5), tetracycline (Wilcoxon, p-value<0.5), and trimethoprim (Wilcoxon, p-value<0.5) classes were consistently reduced. When considering relative abundances, most sampling events were found to experience reductions between 40-60% of ARG classes with S11 (greatest increase) and S17 (greatest decrease) being statistically different (Dunn's Test, p-value<0.05). Sampling events S14 and S16 were the only ones that failed to result in consistent removals when assessing calculated absolute abundances.

In contrast to the flocsed and ozonation processes, BAC5 and BAC10 both tended to be associated with increases in relative and calculated absolute abundances of ARGs. Specifically, BAC5 experienced average increases in 53.3% and 69.2% of antibiotic resistance classes' relative and calculated absolute abundances, respectively, while BAC10 experienced increases of 60.8% and 75.0%, respectively. Further, BAC5 experienced consistent relative abundance enrichment of the multidrug and quinolone antibiotic resistance classes with fosfomycin (Wilcoxon, p-value<0.5) and glycopeptide being consistently reduced. BAC10 experienced consistent relative abundance enrichment of aminocoumarin (Wilcoxon, p-value<0.5), aminoglycoside, beta-lactam, peptide, and tetracycline (Wilcoxon, p-value<0.5) antibiotic resistance classes with reductions of fosfomycin and trimethoprim. When assessing calculated absolute abundances, no antibiotic resistance classes were consistently reduced in either BAC5 or BAC10. Instead, ARGs associated with the aminocoumarin, aminoglycoside, beta-lactam, multi-drug, other, quinolone, sulfonamide, tetracycline, and trimethoprim classes consistently increased in BAC5, while ARGs within the aminocoumarin (Wilcoxon, p-value<0.5), aminoglycoside (Wilcoxon, p-value<0.5), beta-lactam (Wilcoxon, p-value<0.5), fosfomycin, mls, multi-drug (Wilcoxon, p-value<0.5), other, peptide, rifamycin, sulfonamide, and tetracycline (Wilcoxon, p-value<0.5) classes consistently increased in the BAC10. Generally, elevated enrichment of relative abundances of classified ARGs were experienced during S10, S16, and S17 and S5, S10, S14, and S17, for BAC5 and BAC10 respectively. Similar

relationships were seen when considering calculated absolute abundances, where S5, S11, S13, and S17 experienced increases in total ARGs within BAC5 and all except S14 and S16 experienced consistent increases.

Like both BAC process, relative and calculated absolute abundance of total ARGs also tended to increase in both GAC units. On average, antibiotic resistance classes increased in GAC10 57.0% and 53.3% of the time for measures of relative and calculated absolute abundances, respectively. Meanwhile, antibiotic resistance classes, on average, increased in GAC20 57.5% and 50.8% of the time, respectively. GAC10 also consistently increased the relative abundance of aminocoumarin (Wilcoxon, p-value<0.5), glycopeptide, rifamycin (Wilcoxon, p-value<0.5), and tetracycline (Wilcoxon, p-value<0.5) resistance ARGs, while only consistently reducing ARGs classified for quinolone resistance. GAC20 experienced similar relationships as GAC10 for ARGs encoding resistance to aminocoumarin (Wilcoxon, p-value<0.5), glycopeptide, rifamycin, tetracycline (Wilcoxon, p-value<0.5) and quinolone (Wilcoxon, p-value<0.5) antibiotics, in addition to consistent positive selection of other and trimethoprim classifications and removal of sulfonamide (Wilcoxon, p-value<0.5) and mls related ARGs. When assessing calculated absolute abundances, fewer consistent relationships were noted with both GAC contactors experiencing consistent selection of glycopeptide (Wilcoxon, p-value<0.5, only significant for GAC20) and ARGs encoding resistance to tetracyclines, while GAC10 also experienced increases within rifamycin resistance and GAC20 experienced increases within aminocoumarin resistance. GAC20 also noted consistent reductions in calculated absolute abundance of quinolone and sulfonamide ARGs, with no classes being consistently reduced within GAC10.

Following GAC treatment, UV disinfection was, on average, found to successfully reduce both relative and calculated absolute abundances of most annotated ARGs over tested conditions, with notable exceptions. On average, UV treatment found antibiotic resistance classes to decrease 57.1 and 94.6 of the time for measures of relative and calculated absolute abundances, respectively. The former found consistent relative abundance reductions for ARGs associated with aminocoumarin, fosfomycin, glycopeptide, other, phenicol, rifamycin, sulfonamide, tetracycline, and trimethoprim classifications with aminoglycoside, multidrug, peptide, and quinolone (Wilcoxon, p-value<0.5) experiencing consistent positive selection. When considering calculated absolute abundances, every classification of antibiotic resistance was consistently reduced (Wilcoxon, p-value<0.5), with most classes being reduced in 100% of tested conditions, except for elfamycin, which was rarely detected.

When considering the impact of the entire treatment train (Effluent versus Influent), relative abundances were, on average, more commonly enriched over tested conditions, with antibiotic resistant classes experiencing positive selection in 53.1% of comparisons. However, changes in calculated absolute abundances were almost universally reduced, with antibiotic resistant classes experiencing removal 89.8% of the time. Over the course of treatment, the relative abundance of aminocoumarin (Wilcoxon, p-value<0.5), aminoglycoside (Wilcoxon, p-value<0.5), glycopeptide (Wilcoxon, p-value<0.5), multidrug (Wilcoxon, p-value<0.5), other (Wilcoxon, p-value<0.5), rifamycin, and tetracycline were consistently increased while fosfomycin (Wilcoxon, p-value<0.5), phenicol (Wilcoxon, p-value<0.5), sulfonamide, and trimethoprim (Wilcoxon, p-value<0.5) were consistently decreased. When assessing changes in calculated absolute abundances, similarly to final disinfection, all antibiotic resistance classes

except elfamycin and glycopeptide were consistently decreased (Wilcoxon, p-value<0.5) over all tested conditions.

In addition to these treatment processes, comparison were also made between BAC5:BAC10 and GAC10:GAC20 (Figure 3.6). In both comparisons of EBCT, regardless of distinctions between BAC and GAC treatments, higher EBCT resulted in higher average relative and calculated absolute abundances of total ARGs. Specifically, total ARGs were higher at higher EBCT 57% and 63.7% of the time for BAC treatment and 52.1% and 54.9% of the time for GAC treatment based on measures of relative abundance and calculated absolute abundance, respectively. Few ARG classifications experienced consistent selection or removal between tested EBCTs. Within the BACs, fosfomycin, mls, and tetracycline related genes experienced consistently higher relative abundances at higher EBCT while only multidrug ARGs were consistently higher at lower EBCT. When considering calculated absolute abundance changes within the BACs, no ARG class was consistently higher at lower EBCT while aminocoumarin, phenicol, rifamycin, and tetracycline were consistently higher at higher EBCT. Further, only aminoglycoside classified genes were identified as experiencing a consistent shift when compared between GAC10 and GAC20 where they were elevated in the later for both relative and calculated absolute abundance measures.

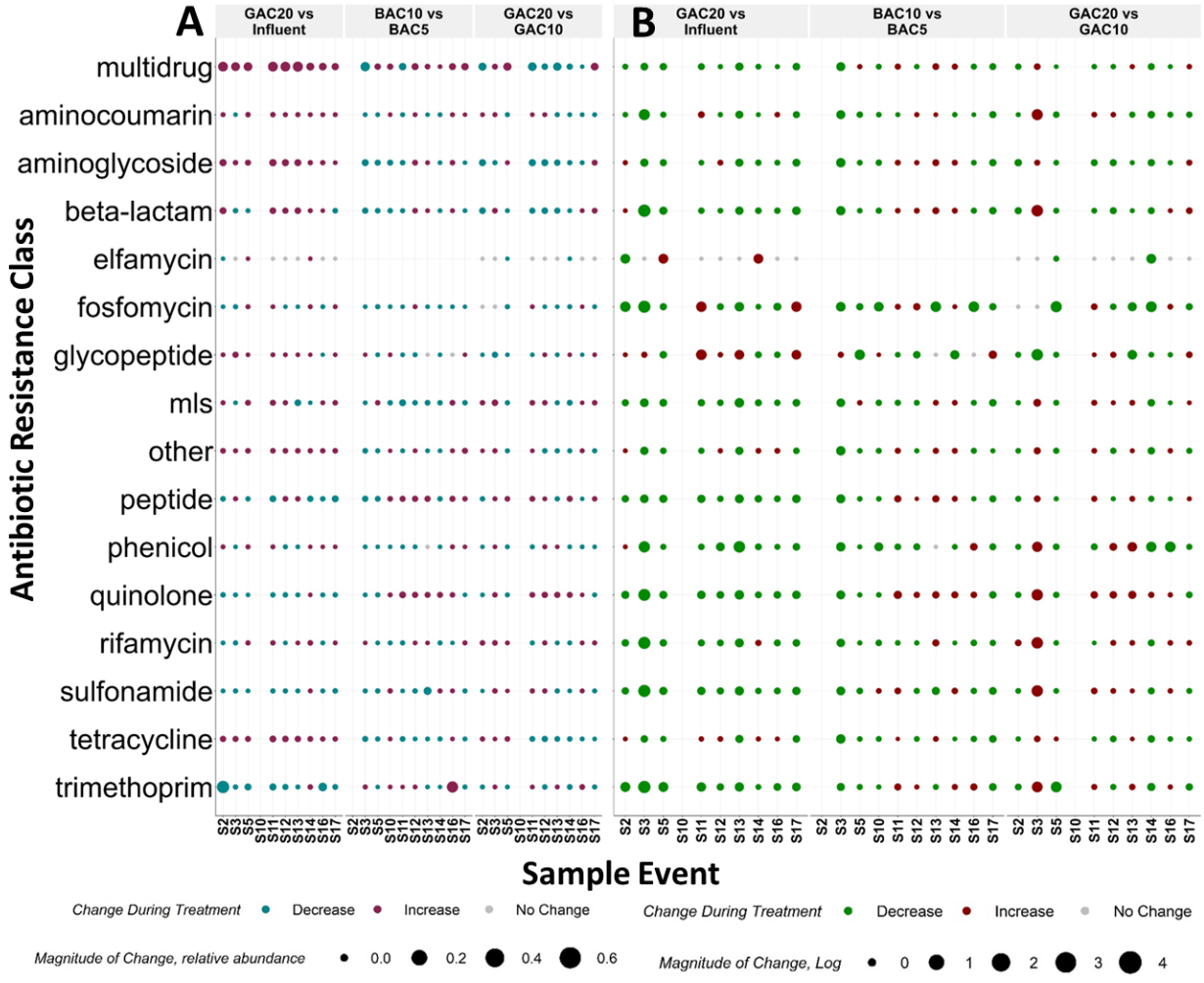


Figure 3.6: Magnitude of change for each antibiotic resistance class between BACs and GACs operated at different EBCTs and BAF for (A) difference in 16S rRNA normalized abundances and (B) log change of the calculated absolute abundances.

In addition to assessing changes between antibiotic resistance classifications within individual treatment processes, temporal variability was also assessed for all comparisons. Kruskal-Wallis statistical testing indicated that changes in classified ARG's relative abundance during ozonation, BAC5, BAC10, GAC10, GAC20, and UV disinfection as well as those between BAC5:BAC10 and GAC10:GAC20 were significant between sampling events (Kruskal-Wallis, p-value<0.05). Further, these significant relationships were expanded to include changes during flocsed treatment, and throughout the entire pilot when considering calculated absolute abundances of classified ARGs. Following this assessment, Dunn's test was applied to identify relationships between groups with complete reporting present in the SI spreadsheet. Dunn's testing grouped sampling events into those that experienced average change in abundances for that comparison, and those that experienced either better or worse than average performance.

SI Figure A.1 and D.6 provides similar relationships for only clinically relevant ARGs as water progresses through each stage of treatment. Here it can be seen that clinically-relevant ARGs associated with glycopeptide and quinolone resistance classes are effectively and completely removed throughout treatment, with the latter only periodically persisting prior to BAC filtration. Only two other resistance classes were associated with any level of clinically-relevant ARGs, including beta-lactam and peptide resistance. Clinically-relevant ARGs associated with beta-lactam resistance experienced consistent relative abundance reductions during the flocsed, BAC5, BAC10 (Wilcoxon, p-value<0.5), GAC10, GAC20 (Wilcoxon, p-value<0.5) treatment processes as well as the entire pilot (Wilcoxon, p-value<0.5) while only experiencing consistent enrichment during ozonation and UV disinfection. Further, no treatment process identified consistent enrichment of calculated absolute abundances with most treatment processes providing consistent reductions including the entire treatment train (Wilcoxon, p-value<0.5). Changes in clinically-relevant ARGs associated with peptide resistance were more variable, with only GAC10 (Wilcoxon, p-value<0.5) providing consistent relative abundance reductions while BAC5, UV disinfection (Wilcoxon, p-value<0.5), and the entire treatment train resulting in enrichment. However, when considering calculated absolute abundances, only BAC5 provided consistent enrichment, while flocsed, ozonation, GAC20, UV disinfection, and the entire treatment train (Wilcoxon, p-value<0.5) provided consistent removal. Unlike bulk ARG measures, clinically-relevant ARGs associated with beta-lactam and peptide resistances were also more commonly higher in shorter EBCT for both BAC and GAC treatments. This was true both for relative and calculated absolute abundances. Kruskal-Wallis statistical testing also failed to indicate any significant differences between sampling events at each treatment process.

Identification and Comparison of Core ARGs

Throughout all stages of treatment and sampling events, 544 unique ARGs were detected. Of these 544 ARGs, ozone effluent accounted for the highest number detected (373), while GAC20 accounted for the lowest number detected (288). Further, influent samples contained 349 of the 544, while flocsed, BAC5, BAC10, GAC10, and UV effluents contained 305, 325, 307, 300, and 310 detected genes, respectively. When accounting for average unique genes per sampling event, these total detected gene counts dropped to 154, 146, 161, 146, 144, 139, 126,

and 149 for influent, flocsed, ozone, BAC5, BAC10, GAC10, GAC20, and UV effluent samples, respectively.

These unique ARGs make up the ‘resistome’, which can be further described by core, and discriminatory components. The former was considered any gene that was consistently detected in all sampling events and over all tested conditions for each treatment effluent. These ARGs are shown in Figure 3.7 and summarized in Table 3.1, with counts provided in Table 3.2 along with the average number of unique ARGs associated with each antibiotic resistance classification. On average, the core component of ARGs accounted for roughly 36% of unique ARGs between all sampling locations - though this was influenced by the relatively low amount of core genes associated with ozonation (17/161) and GAC20 (19/126), with all other treatment processes having core genes making up between 40-50% of the average identified genes. Overall, ARGs associated with the multidrug and peptide resistance classifications accounted for the greatest number of core ARGs. When considering clinically-relevant ARGs, only 5 were identified as being core to any treatment process including 3 peptide resistance ARGs (MCR-2, MCR-4, and MCR-5) and 2 beta-lactam resistance genes (OXA-5 and TEM-159). MCR-5 was found to be the only clinically-relevant ARG core to multiple treatment processes – specifically influent, BAC5, and UV effluents - while OXA-5 (influent), MCR-2 (flocsed), MCR-4 (BAC5), and TEM-159 (BAC5) were only core to individual processes. No clinically-relevant ARGs were found to be core to ozone, BAC10, GAC10, or GAC20.

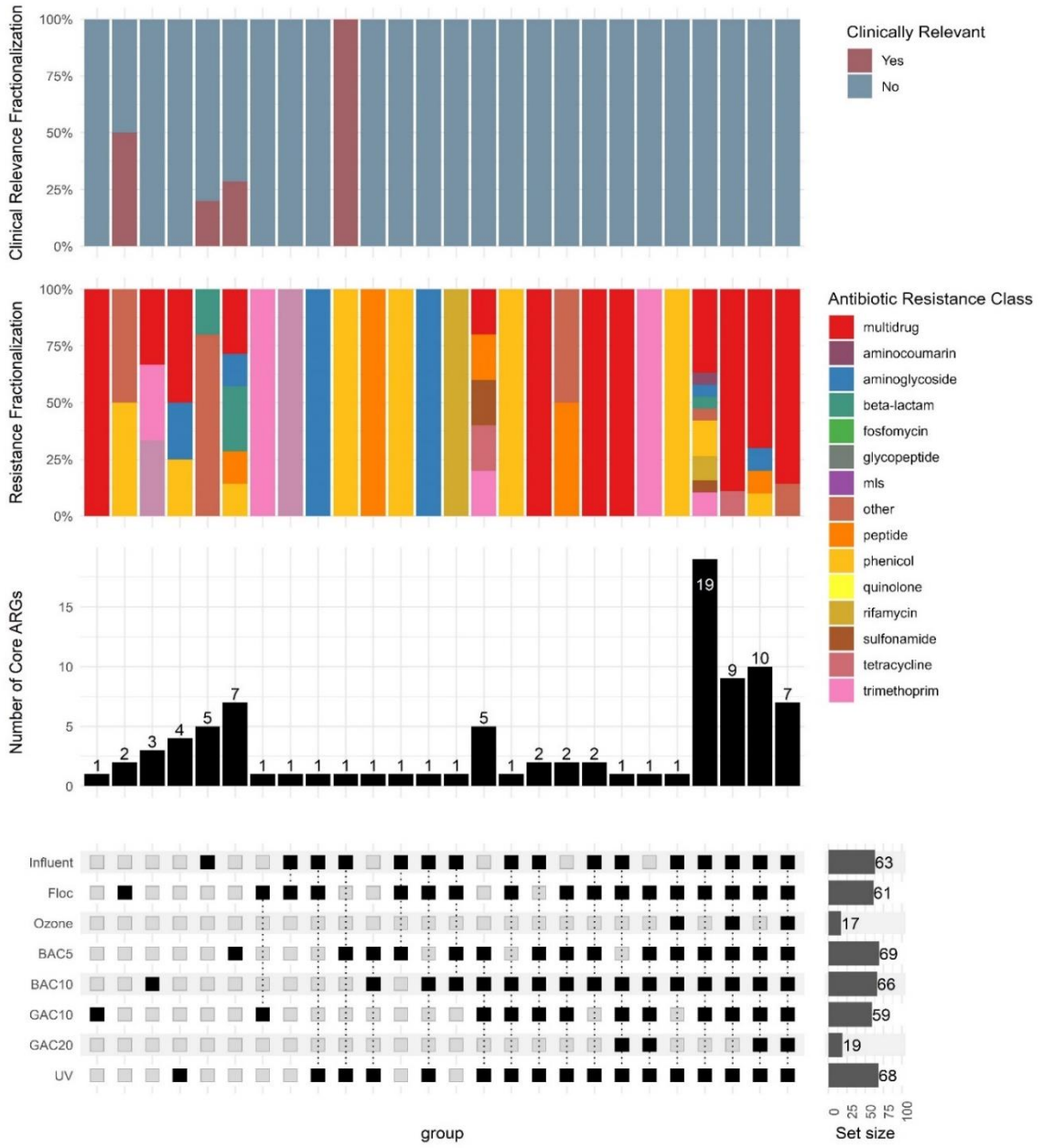


Figure 3.7: Provides a breakdown of all core ARGs (ARGs found at every single sampling event) identified throughout treatment and the relationships those core ARGs have with multiple treatment processes. ARGs are also classified by their clinical relevance and resistance classification. Set size provides the total number of core ARGs for the associated treatment process.

Table 3.1: Core ARGs identified at each sampling location, by antibiotic resistance class. (*) Indicates a clinically relevant ARG.

ARG Classification	Treatment Location							
	Influent	**FlocSed**	**Ozone**	**BAC5**	**BAC10**	**GAC10**	**GAC20**	**UV**
aminocoumarin	novA	novA	N/A	novA	novA	novA	N/A	novA
aminoglycoside	AAC(6')-Ib7, acrD, amrA, amrB	AAC(6')-Ib7, acrD, amrA, amrB	N/A	aadA14, acrD, amrB	acrD, amrA, amrB	acrD, amrB	amrB	AAC(6')-Ib7, aadA5, acrD, amrA, amrB
beta-lactam	Burkholderia-pseudomallei-Omp38, OXA-5*	Burkholderia-pseudomallei-Omp38	N/A	Burkholderia-pseudomallei-Omp38, LRA-1, TEM-159*	Burkholderia-pseudomallei-Omp38	Burkholderia-pseudomallei-Omp38	N/A	Burkholderia-pseudomallei-Omp38
mls	macB, mphA, mphD, Mrx, msrB, msrE	ErmF, LpeB, macB, msrB	macB	LpeB, macB, msrB	LpeB, macB, msrB	LpeB, macB, msrB	macB	LpeB, macB, msrB
multidrug	acrB, acrF, adeF, adeJ, AxyY, ceoA, ceoB, cmeB, emrA, emrB, mdsB, mdtB, mdtC, mdtF, MexB, MexC, MexD, MexF, mexI, mexK, mexN, mexQ, mexW, mexY, msbA, mtrD, MuxB, MuxC, opcM, oqxB, sav1866, smeB, smeE	acrB, acrF, adeF, adeJ, AxyY, ceoA, ceoB, cmeB, emrB, mdsB, mdtB, mdtC, mdtF, MexB, MexD, MexF, mexI, mexK, mexN, mexQ, mexW, mexY, msbA, mtrD, MuxB, MuxC, opcM, oqxB, sav1866, smeB, smeE	adeF, mdsB, mdtC, MexD, mexK, mexN, mexQ, mexW, msbA, MuxB, oqxB, sav1866, smeB, smeE	acrB, acrF, adeF, adeG, adeJ, AxyY, ceoB, cmeB, emrA, emrB, mdsB, mdtB, mdtC, mdtF, MexB, MexC, MexD, MexF, mexI, mexK, mexN, mexQ, mexW, mexY, msbA, mtrD, MuxB, MuxC, opcM, oprA, OprN, oqxB, sav1866, smeB, smeE	acrB, acrF, adeF, adeJ, AxyY, ceoA, ceoB, cmeB, emrA, emrB, mdsB, mdsC, mdtB, mdtC, mdtF, MexB, MexC, MexD, MexF, mexI, mexK, mexN, mexQ, mexW, mexY, msbA, mtrD, MuxB, MuxC, opcM, oprA, oqxB, sav1866, smeB, smeE	acrB, acrF, adeF, adeJ, AxyY, ceoA, ceoB, cmeB, efrB, emrA, mdsB, mdtB, mdtC, mdtF, MexB, MexC, MexD, MexF, mexI, mexK, mexN, mexQ, mexW, mexY, msbA, mtrD, MuxB, MuxC, oprA, oqxB, sav1866, smeB, smeE	acrB, adeF, ceoA, ceoB, mdsB, mdtB, mdtC, MexB, MexD, MexF, mexY, msbA, MuxB, MuxC	acrB, acrF, adeF, adeJ, AxyY, ceoA, ceoB, cmeB, emrA, emrB, emrK, mdsB, mdtB, mdtC, mdtF, MexB, MexC, MexD, MexF, mexI, mexK, mexN, mexQ, mexW, mexY, msbA, mtrD, MuxB, MuxC, opcM, oprA, OprM, oqxB, sav1866, smeB, smeE
other	TriC	Bifidobacteria-intrinsic-ileS-conferring-resistance-to-mupirocin, TriC		Bifidobacteria-intrinsic-ileS-conferring-resistance-to-mupirocin, farA, farB, optrA, TriC	Bifidobacteria-intrinsic-ileS-conferring-resistance-to-mupirocin, farA, optrA, TriC	Bifidobacteria-intrinsic-ileS-conferring-resistance-to-mupirocin, optrA, TriC	TriC	Bifidobacteria-intrinsic-ileS-conferring-resistance-to-mupirocin, farA, optrA, TriC
peptide	arnA, bacA, bcrA, MCR-5*, PmrC, pmrE, rosA, rosB	arnA, bacA, bcrA, MCR-2*, PmrC, pmrE, rosA, rosB	pmrE	arnA, bacA, MCR-4*, MCR-5*, PmrC, pmrE, rosA, rosB	arnA, bacA, bcrA, pmrE, rosA, rosB	arnA, bacA, bcrA, rosA, rosB	rosB	arnA, bacA, bcr-1, bcrA, MCR-5*, pmrE, rosA, rosB
quinolone	qacH, qepA, QepA3	qacH, qepA, QepA3	N/A	qacH, qepA, QepA3	qacH, qepA, QepA3	qepA, QepA3	N/A	qepA, QepA3
rifamycin	rphB	rphB	N/A	rphA, rphB	rphA, rphB	rphA, rphB	N/A	rphA, rphB
sulfonamide	sul1	sul1	sul1	sul1, sul2	sul1, sul2	sul1, sul2	N/A	sul1, sul2
tetracycline	adeB, tet(41)	adeB, tet(41), tet(G), tetA(48)	N/A	adeB, tet(41), tetA(48), tetT	adeB, tet(41), tetA(48), tetB(60), tetT	adeB, tet(41), tet(G), tetA(48), tetT	tetA(48)	adeB, tet(41), tetA(48), tetT
trimethoprim	dfrB6	dfrB6	N/A	N/A	dfrA3	N/A	N/A	N/A

Figure 3.7 provides a graphical representation of the shared, core ARGs between all treatment process comparisons while SI Figure A.1 provides the same comparison for all unique ARGs identified during testing. Supplemental graphics are also provided for both comparisons to highlight selected comparisons of interest, SI Figure D.8 and D.9. Assessment of the core genes present prior to BAF (i.e. influent, flocsed, and ozonation) shows that a striking majority of core genes are shared between the influent and flocsed process (38), with only 8 and 6 unique core genes identified at each treatment process, respectively. The ozonation process is also identified as not containing a single core ARG that was not also core to both the influent and flocsed systems where 17 additional genes were identified. Ozone similarly contains no unique, core ARGs when compared to all treatment process, and only fails to share core ARGs with GAC10 and GAC20. The UV process contains slightly more unique, core ARGs (13) when compared to the influent (8) and ozone (0) processes, but still shares most of its core ARGs with influent (55) and ozone (17). When comparing influent samples to both BACs and GACs operated at multiple EBCTs, the majority of core ARGs are again shared with unique, core ARGs only making up between 1.4% (GAC20) to 19.5% (BAC5) of the core resistome. Comparing BAC5 to BAC10 indicated that the lower EBCT with more unique, core ARGs (9 vs 6) which was similar to the comparison of GACs where GAC10 also contained more unique, core ARGs (40 vs 0). The identified relationship of lower EBCT resulting in more unique, core ARGs was also found to be true when comparing all unique ARGs, where BAC5 effluent was found to contain 77 unique ARGs compared to BAC10's 59 and GAC10's 61 vs GAC20's 49. It is also worth noting that GAC20 experienced a lot fewer core ARGs, when compared to other process effluents, but similar levels of total unique ARGs.

Table 3.2: Number of Core ARGs at each sampling location (Average number of unique genes at each location).

ARG Classification	Influent	FlocSed	Ozone	BAC5	BAC10	GAC10	GAC20	UV
aminocoumarin	1 (1)	1 (1)	0 (1)	1 (1)	1 (1)	1 (1)	0 (1)	1 (1)
aminoglycoside	4 (13)	4 (15)	0 (14)	3 (16)	3 (16)	2 (13)	1 (13)	5 (16)
beta-lactam	2 (18)	1 (15)	0 (21)	3 (16)	1 (15)	1 (14)	0 (10)	1 (15)
elfamycin	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
fosfomycin	0 (1)	0 (2)	0 (2)	0 (1)	0 (2)	0 (1)	0 (1)	0 (1)
glycopeptide	0 (1)	0 (0)	0 (1)	0 (1)	0 (2)	0 (2)	0 (2)	0 (2)
mls	6 (18)	4 (16)	1 (16)	3 (10)	3 (11)	3 (11)	1 (9)	3 (11)
multidrug	33 (52)	31 (49)	14 (55)	35 (56)	35 (54)	33 (51)	14 (46)	36 (56)
other	1 (5)	2 (5)	0 (6)	5 (7)	4 (6)	3 (6)	1 (5)	4 (7)
peptide	8 (12)	8 (12)	1 (11)	8 (11)	6 (11)	5 (11)	1 (10)	8 (13)
phenicol	0 (2)	0 (3)	0 (3)	0 (2)	0 (2)	0 (1)	0 (2)	0 (1)
quinolone	3 (5)	3 (5)	0 (5)	3 (5)	3 (5)	2 (5)	0 (3)	2 (5)
rifamycin	1 (5)	1 (4)	0 (5)	2 (4)	2 (5)	2 (5)	0 (5)	2 (5)
sulfonamide	1 (3)	1 (3)	1 (2)	2 (2)	2 (2)	2 (3)	0 (3)	2 (3)
tetracycline	2 (13)	4 (12)	0 (15)	4 (12)	5 (12)	5 (13)	1 (12)	4 (13)
trimethoprim	1 (5)	1 (5)	0 (4)	0 (2)	1 (2)	0 (3)	0 (3)	0 (2)
Total	63 (154)	61 (146)	17 (161)	69 (146)	66 (144)	59 (139)	19 (126)	68 (149)

When comparing all treatment processes together (Figure 3.7) it can again be seen that BAC5 was characterized by the greatest number of unique, core ARGs (7) followed by influent (5), UV (4), BAC10 (3), flocsed (2), and GAC10 (1). Neither ozone nor GAC20 displayed any unique, core ARGs with both sharing their core ARGs with multiple treatment processes. Specifically, ozonation shared all its core ARGs with the entire pilot (7), entire pilot with the exception of GAC20 (9), and the entire pilot with the exception of both GAC units (1), while GAC20 shared all its core ARGs with the entire pilot (7), entire pilot sans ozone (10), entire pilot sans influent and ozone (1), and entire pilot sans ozone and BAC5 (1). Further, of the 90 unique, core ARGs identified throughout all treatment processes, most were associated with greater than 6 treatment process (48), or singular treatment processes (22). The remaining unique, core ARGs (20) were associated with between 2 and 5 treatment processes with most relationships involving either influent|flocsed and/or BAC5|BAC10.

Identification of Discriminatorily Abundant ARGs

LefSe analysis was conducted on the 16S rRNA gene normalized abundances of each sample's ARG composition to identify abundantly discriminatory ARGs between each treatment process, referred to as discriminatory genes. When analyzing the entire dataset together, statistically significant, discriminatory ARGs were identified for each treatment process (Figure 3.8). Influent, flocsed effluent, ozone effluent, and UV effluent were all associated with a single discriminatory ARG – pmrE:peptide, msrE:MLS, Mrx:MLS, and oprA:multidrug, respectively. BAC effluent was distinguishable with multiple ARGs including mtrD:multidrug, abeS:multidrug , aadA14:aminoglycoside, AxyX:multidrug similarly to GAC effluent which was represented by mdsB:multidrug, and AAC(3)-IV:aminoglycoside.

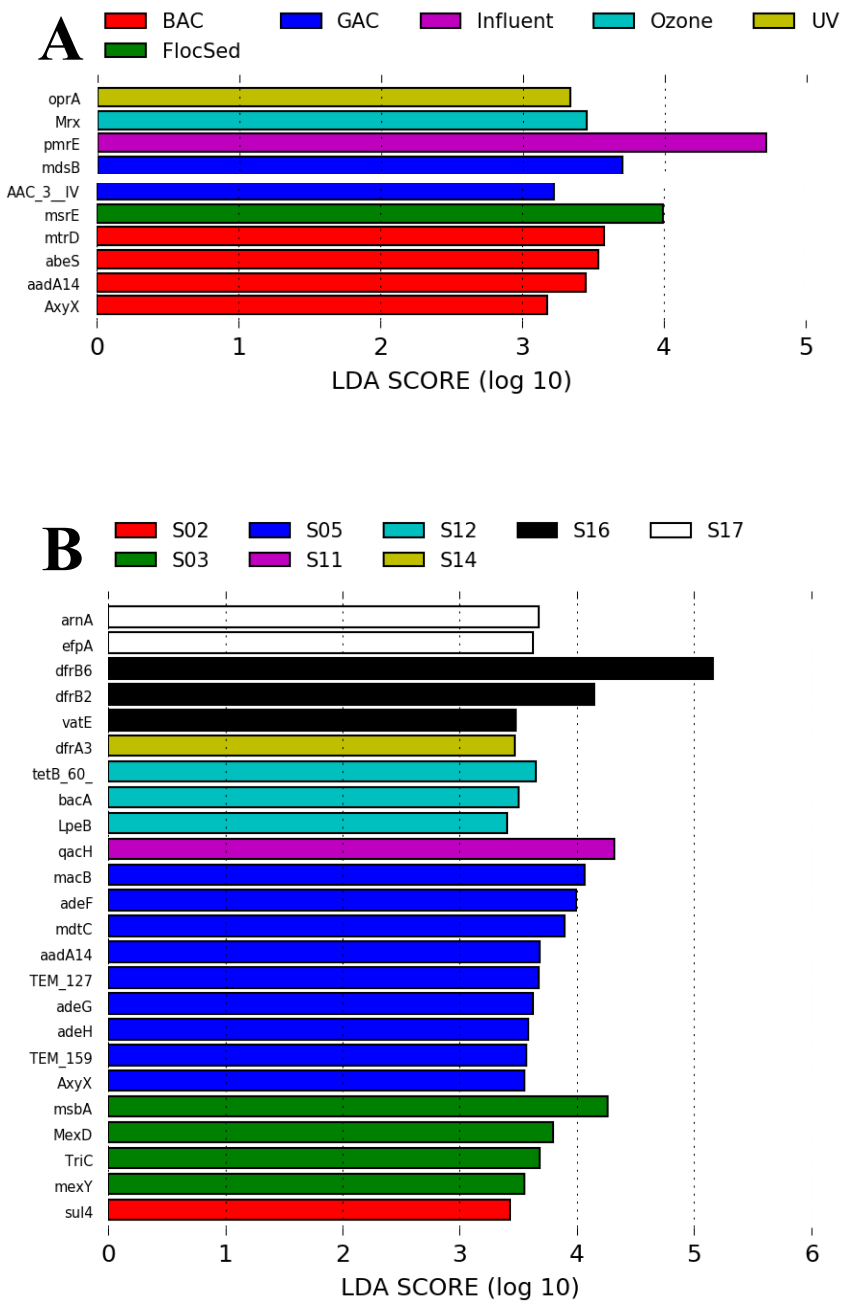


Figure 3.8: Discriminatorily abundant ARGs (determined via LefSe analysis) for (A) treatment processes and (B) sampling events based on 16S rRNA gene normalized relative abundances.

Significantly more discriminatory ARGs were identified when comparing specific treatment processes, instead of all treatment processes at once. A Full compilation of this analysis is reported in Appendix B, with key trends noted here. LefSe analysis identified only 4 ARGs that were discriminatory between influent (3) and the flocsed (1) process while 33

discriminatory genes were identified between the flocsed effluent (18) and ozone effluent (15). Comparisons between ozone effluent and both BAC trains found more total discriminatory ARGs at the longer EBCT (29 vs 23) with both BAC5 (14) and BAC10 (16) containing more discriminatory ARGs than the ozone process (9 and 13 respectively). Comparisons between the two tested EBCT trains interestingly found more discriminatory ARGs associated with the shorter EBCT (BAC5:GAC10 = 42 vs BAC10:GAC20 = 29), though this was influenced by the lack of discriminatory ARGs identified within GAC20 (6) compared to GAC10 (30) when analyzed against their corresponding contactor influents. Further, GAC20 (3) also identified fewer genes when directly compared to GAC10 (14). Conversely, analysis of the BAC related comparisons instead found that longer EBCTs resulted in elevated counts of discriminatory ARGs. Specifically, BAC10 (23) vs BAC5 (12) when compared to their respective GAC contactors, BAC10 (9) vs BAC5(6) when compared to each other, and in the comparison to ozone effluent. UV disinfection resulted in relatively few (17) discriminatory ARGs between it (16) and GAC20 (1). However, when compared to samples collected at the inlet of the pilot (32), UV effluent (47) had significantly more discriminatory ARGs. GAC20 (37) experienced similar increased in discriminatory ARGs when compared to pilot influent (41).

Outside of discriminatory ARGs being linked to differences between treatment processes, LefSe analysis also identified several ARGs associated with each sampling event (Figure 3.8). Most sampling events were linked to 1 to 4 discriminatory ARGs while 9 discriminatory ARGs were associated with S05.

DISCUSSION

This study provides a comprehensive assessment of the fate and composition of ARGs throughout multiple stages of a carbon-based advanced treatment train intended for indirect potable reuse. Samples were collected over a 20-month period that included multiple changes in conditions reflective of reasonable operation and typical water quality variation before being subjected to metagenomic sequencing. ARGs were assessed as a comprehensive resistome, individual core and discriminatory genes, and total ARGs annotated to CARD to better understand the complex nature of AMR within a major barrier for environmental dissemination. Additionally, total and clinically relevant ARGs were assessed in terms of their relative and calculated absolute abundances to better understand selection dynamics within cells and total ARG loadings. Ultimately, both measures of AMR are justified, and additional research needs to be completed to better understand implications for public health as each measures provides unique pros and cons.

For instance, relative abundance is useful for understanding how various processes might select for or against carriage of ARGs by bacteria[37–39], while calculated absolute abundance is useful for determining total removal of ARGs at each unit process[4,40]. Aside from not being able to translate 16S rRNA gene copy numbers into cell numbers, the normalization of ARGs to 16S rRNA and reporting concentrations in terms of relative abundance has some additional limitations, including: 1) detection of DNA sequences does not provide direct information about cell viability, i.e., DNA may persist outside of the cell or in a nonviable cell as the result of ozone oxidative damage or UV irradiation 2) shorter basepair Illumina reads may still be intact and have identifiable reads and 3) PCR amplification is hindered by damaged DNA[41,42]. Further, relative abundance does not account for potential relative changes in the DNA composition from extracellular and intracellular DNA contributions depending on extraction methods or sample

treatments. Additionally, relative abundance can mask an increase or decrease in absolute ARG copy numbers resulting from growth of a resistant population thus calculated absolute abundance data can be used for this purpose.

Overall, the fate of ARGs throughout the pilot and individual treatment processes were heavily influenced by measures of relative or calculated absolute abundance. Processes that decreased calculated absolute abundances through sheer reductions in the 16s rRNA gene marker for total bacterial counts (e.g., ozonation and UV disinfection) also resulted in relative abundance increases, further complicating the role of treatments in ARG mitigation. Throughout the entire treatment train, relative abundances experienced a roughly 60% increase. However, it is worth noting that both the influent and effluent abundances are notably lower than other studies looking at domestic wastewater effluents[4], international wastewater influents[11], and various forms of reuse waters[19,43] indicating that treatment does still provide an effective barrier to ARG proliferation in line with other wastewater derived systems. Further, calculated absolute abundances of ARGs experienced an average reduction of 2-logs.

Intermediate processes also provided notable effects to ARG proliferation and mitigation with the BAC filters generally providing positive selection[44–46] for both relative and absolute measures while flocsed provided the opposite[47–49]. GAC contactors provided less straight forward interactions with ARGs[50,51] where their relative abundances typically decreased and calculated absolute abundances typically increased. In all cases, clinically relevant ARGs made up an extremely small proportion of the total ARG loading and often had their calculated absolute abundances decreased outside of BAC filtration.

BAC filtration was ultimately identified as an extremely important intermediate treatment process as it was one of the primary processes that increased ARG abundances. It is hypothesized that BAC filters and later stage GAC contactors, post-exhaustion, can act as reservoirs for ARGs[44]. It is possible that adsorptive media, like what is found in BAC filters, facilitate interactions between microorganisms and chemical compounds (including antibiotics and other CECs) both in the term of contact times and chemical concentrations when compared to bulk water interactions. If true, BACs could provide an environment that is conducive to selective enrichment that are often absence in non-attached growth process. However, more targeted research into these interactions is needed.

In addition to changing the abundances of individual ARGs and ARGs grouped by resistance class, each treatment process outside of flocsed also resulted in significant differences to the resistome composition. Interestingly, for both BAC and GAC treatments the EBCTs tested failed to produce a significant change when compared within the treatment process but did produce significant differences when compared between processes. This isn't necessary surprising for a new or recently regenerated GAC contactor but considering several sampling events were collected after GAC exhaustion, where essentially physical removal mechanisms begin to shift toward biological, the resistome still seemed to be shaped by changes in nutrient profiles and differences between bacterial colonization and succession. It is also important to acknowledge that this study, and others, have found that factors that influence changes in taxonomy will also alter the resistome[52–56]. With that in mind, changes to the resistomes composition within the BAC filters were also found to be related to the addition of supplemental nutrients, polymer additions, and changes to chlorine quenching agents which were all identified to have varying degrees of impact on the microbial community composition (Chapter 5).

However, whether these operational changes directly resulted in changes to ARG composition or indirectly through changes in microbial composition remains unclear. Similarly, resistome composition was also related to pilot runtime, indicating that a temporal component to resistome formation was also present within the BAC filters. Besides the biologically active filter, ozonation was the only treatment process that experienced inter-process variation that was associated with an operational change. Specifically, the calculated absolute abundance measure of the resistome were related to ozone dosing with those subjected to higher ozone doses clustering together. Again, these factors are likely partially related to changes in bacterial numbers and community makeup where elevated ozone dosing provided greater changes in total bacterial gene markers and has been identified to alter ARG concentrations[57–60].

When considering the fate of specific classes of ARGs, it was generally seen that the flocsed process provided effective reduction with only aminoglycoside experiencing consistent relative abundance increased. Flocsed was also more effective at reducing calculated absolute abundances of ARG likely due to increase solid liquid separation and associated removal of bacterial numbers. The flocsed process was also found to be generally consistent in its operation and removal of ARGs, failing to find statistical differences between different sampling events. Conversely, ozonation provided extremely variable changes to ARG relative abundances with most classifications falling close to a 50:50 split between enrichment and removal. However, as to be expected with disinfection processes, ozonation provided sufficient reductions in total bacterial numbers and therefore also effectively reduced the total loading of ARGs following treatment. Further, S14 and S16 where the only sampling events that failed to experience consistent reductions of calculate absolute abundances with S14 being subjected to the lowest ozone dose while S16 experienced elevated turbidity and nitrite concentrations which acts as an additional ozone sink[61,62]. This indicates that appropriate ozone dosing and ratios of ozone generation to oxidative sinks is paramount to effectively limiting the total ARG loading.

Even though the resistome based analysis failed to find significant differences between BAC filters and GAC contactors at the tested EBCTs, differences were found when assessing the change in classified ARGs during treatment. Specifically, higher EBCTs typically resulted in more consistent changes in ARG classes over multiple sampling events speaking to an increased volatility with higher EBCTs. However, since the changes within both BAC and GAC processes typically resulted in increased abundances of ARGs, higher consistency also resulted in more consistent enrichment for total ARGs. This is possibly due to the increased facilitation of microbe to compound, and microbe to microbe interactions at lower flow rates. Similarly, to how reductions in total measures of bacterial populations can overshadow relative abundance increase during disinfection processes, the opposite was found to be true with calculated absolute abundance measures during BAC treatment where bacterial numbers rebounding following ozonation resulting is elevated enrichment of ARG classes. Since similar proliferation of bacterial population markers did not occur during GAC treatment, calculated absolute abundances experienced closer to a 50:50 split between enrichment and removal even though relative abundance measures experienced similar levels of enrichment during BAC filtration. In both BAC and GAC treatments, a comparison of EBCT indicated that in addition to higher EBCT resulting in more consistent changes over time, they also resulted in higher average relative and calculated absolute abundances of classified ARGs.

Similarly to ozonation, UV treatment was found to be extremely effective at reducing calculated absolute abundances of all ARG classes. Unlike ozonation, UV disinfection was also found to be effective at consistently removing a majority of ARG classes' relative abundances. However, genes associated with a few antibiotic resistance classes were significantly, and consistently enriched, overshadowing the reduction experienced within the majority of ARG classes and ultimately resulting in cumulative increases in total ARG relative abundances. These classes include genes associated with aminoglycoside, multidrug, peptide, and quinolone resistances indicating that special considerations are required for certain resistant groups.

Over the entirety of the pilot train, classified bulk ARG's absolutely abundances were ubiquitously removed, while relative abundances were found to experience enrichment more often (enrichment occurred in 53.1% of the time in classified ARG groupings) than they were reduced, largely due to enrichment occurring during BAC and GAC treatment. Specifically, the classes related to aminocoumarin, aminoglycoside, glycopeptide, multidrug, other, rifamycin, and tetracycline resistance we consistently selected for during tested conditions while fosfomycin, phenicol, sulfonamide, and trimethoprim resistance classes were consistently reduced. The dichotomy between these classes of ARGs speaks to the value that metagenomics provides in assessing something as complex as AMR and the difficulty in selecting appropriate indicator genes to accurate monitor trends in AMR proliferation and mitigation since the results can be heavily determined by the selected genes of interest.

Considerations are further complicated when factoring in the differences between trends during treatments with total ARGs and those of clinical relevance. During this study, clinically relevant genes made up a much smaller proportion of total ARGs and were generally better removed during treatment process with fewer persisting throughout treatment. Only clinically relevant ARGs associated the peptide resistance classification experienced consistent relative abundance increases, largely due to the UV disinfection process, while all clinically relevant ARGs were consistently reduced in absolute abundance. Unlike total ARG measures, clinically relevant genes associated with beta-lactam and peptide resistances were found more commonly at higher concentrations in shorter EBCT for both BAC and GAC treatments, regardless of distinctions between relative and calculated absolute abundances. Further, BAC filtration and GAC contacting didn't result in relative abundance selection for clinically relevant ARGs as commonly as they did for bulk ARG comparisons.

When assessing the individual genes associated with the resistome, it was found that there were significantly more unique genes identified over multiple samplings than the average number of genes for any given sampling event. This indicates that there was inherent, temporal variability associated with the ARGs making up the resistome which was similarly found in several studies on wastewater derived samples[40,63] and soil based resistomes[64]. However, throughout this variability a core group of genes was still identified across all tested conditions indicating that the resistome is made up of both core and transient genes with certain genes being discriminatory between specific treatment processes.

Overall, every tested treatment process, except ozonation and GAC20, experienced between 59 to 69 core ARGs out of between 139 to 154 average, unique genes. However, even though the magnitude of the core resistome was consistent, the specific gene profiles moving through the treatment trains were ever changing which supports the continued conclusion that treatment processes impact the formation of the resistome[4,65–68]. Ozone effluent, which

experienced the lowest number of core ARGs, interestingly had the higher number of average, unique ARGs (161) indicating both a high level of volatility associated with the ozonation process and responsiveness of the resistome to changing operational conditions. Specifically, the ozone process was subjected to a wide range of ozone doses and targeted ozone:TOC ratios which likely resulted in the low number of core ARGs found throughout all sampling events. GAC20 experienced similarly low numbers of core ARGs as well as the lowest number of average, unique ARGs (126) both of which were much lower than what was experienced in GAC10 seeming to indicate that longer EBCTs during GAC contacting can result in fewer core and average, unique ARGs. Similarly, BAC5 and GAC10 both experienced higher numbers of core and average, unique ARGs when compared to their respective BAC10 and GAC20 counterparts supporting the hypothesis that lower EBCTs may result in a more diverse resistome. Additionally, both disinfection processes (ozonation and UV disinfection) were the only treatment process that resulted in higher average, unique ARGs than the treatment process directly upstream.

Core ARGs were also heavily shared as all treatment processes, except BAC5, experienced more than 90% of their core ARGs being associated with at least one other treatment process. Shared genes were also linked to upstream treatments where all but 6 genes were found to be core in either the pilot influent and/or flocsed effluent showing the impact of influent resistome characteristics on preceding processes. Conversely, BAC5 experienced the greatest number of total core ARGs and core ARGs unique to a single treatment. This supports the hypothesis that BAC filters, especially at lower EBCT, may act as a reservoir for ARG selection as chemical compounds (both adsorbed to the media and present in the bulk water) come in continuous contact with complex, microbial communities. Proximity to the ozonation process, where average, unique ARGs were found to be in highest concentrations, and increased mass loading of contaminants at lower EBCT likely contribute to this phenomenon more so than during GAC contacting or BAC filtering at higher EBCT. It is also important to note that few ARGs identified as being clinically relevant were found to be core to any treatment process, with only MCR-5 being consistently found in the UV effluent.

When assessing discriminatorily abundant ARGs, all treatment processes were found to be distinguishable by at least a single ARG over all tested conditions. Interestingly, BAC filtration and GAC contacting were found to have the most discriminatory ARGs, speaking to the unique nature of the resistome composition during BAF treatment. Significantly more discriminatory ARGs were found when comparing between specific treatment processes instead of the treatment train in its' entirety adding to the evidence that certain treatment processes do impact and shape the resistome. The flocsed process was found to not be extremely different from that of the influent process with comparatively few discriminatory ARGs identified. Interestingly, more discriminatory genes were found between BAC:GAC comparisons than between BAC5:BAC10 or GAC10:GAC20 comparisons indicating that even late-stage GAC contactors (essentially operating as BAC filters) were still distinguishable from their preceding BAC treatments. These differences are theorized to be related to changes in operation (proximity to ozonation, and prolonged operation without backwashing within the GAC contactors), differences in substrate concentrations, and variable microbial communities. There were also significant numbers of discriminatory ARGs between the front (influent) and end (GAC20 and UV effluent) of the treatment train providing additional evidence for a consistent change in the resistome throughout the entire treatment train.

Additional discriminatory ARGs were also found between sampling events, demonstrating that the resistome is also susceptible to temporal variation. This variation is likely impacted by the changing operational conditions, water quality parameters, and microbial succession occurring throughout operation. Therefore, temporal understanding of resistome dynamics should not be ignored in future longitudinal studies even if spatial factors were found to be dominant in the core analysis.

ACKNOWLEDGEMENTS

The authors would like to thank the operators and staff at Hampton Road Sanitation District's York River Treatment Plant and the SWIFT Research Center. Funding for this effort was provided in part by Water Research Foundation Unsolicited Award U1R16 (PI Pruden) and US Bureau of Reclamation grant R21AC10162 (PI Pruden) 09/01/2020-08/31/2023 leveraged with additional financial and in-kind support from the Hampton Roads Sanitation District. Special thanks to Pruden Lab members for assistance in the laboratory.

REFERENCES

1. Sano, D., Amarasiri, M., Hata, A., Watanabe, T. & Katayama, H. Risk management of viral infectious diseases in wastewater reclamation and reuse: Review. *Environ. Int.* **91**, 220–229 (2016).

2. Harwood, V. J. *et al.* Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. *Appl. Environ. Microbiol.* **71**, 3163–3170 (2005).

3. Aarestrup, F. M. & Woolhouse, M. E. J. Using sewage for surveillance of antimicrobial resistance. *Science (80-.).* **367**, 630–632 (2020).

4. Majeed, H. J. *et al.* Evaluation of Metagenomic-Enabled Antibiotic Resistance Surveillance at a Conventional Wastewater Treatment Plant. *Front. Microbiol.* **12**, (2021).

5. Auerbach, E. A., Seyfried, E. E. & McMahon, K. D. Tetracycline resistance genes in activated sludge wastewater treatment plants. *Water Res.* **41**, 1143–1151 (2007).

6. Chen, H. & Zhang, M. Occurrence and removal of antibiotic resistance genes in municipal wastewater and rural domestic sewage treatment systems in eastern China. *Environ. Int.* **55**, 9–14 (2013).

7. Gao, P., Munir, M. & Xagoraraki, I. Correlation of tetracycline and sulfonamide antibiotics with corresponding resistance genes and resistant bacteria in a conventional municipal wastewater treatment plant. *Sci. Total Environ.* **421**–**422**, 173–183 (2012).

8. Wright, G. D. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat. Rev. Microbiol. 2007 53* **5**, 175–186 (2007).

9. Yadav, S. & Kapley, A. Exploration of activated sludge resistome using metagenomics. *Sci. Total Environ.* **692**, 1155–1164 (2019).

10. Bengtsson-Palme, J. *et al.* Elucidating selection processes for antibiotic resistance in

sewage treatment plants using metagenomics. *Sci. Total Environ.* **572**, 697–712 (2016).

11. Prieto Riquelme, M. V. *et al.* Demonstrating a Comprehensive Wastewater-Based Surveillance Approach That Differentiates Globally Sourced Resistomes. *Environ. Sci. Technol.* (2022) doi:10.1021/acs.est.1c08673.

12. Liu, Z. *et al.* Metagenomic and metatranscriptomic analyses reveal activity and hosts of antibiotic resistance genes in activated sludge. *Environ. Int.* **129**, 208–220 (2019).

13. Hendriksen, R. S. *et al.* Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun. 2019 101* **10**, 1–12 (2019).

14. Baral, D. *et al.* Tracking the Sources of Antibiotic Resistance Genes in an Urban Stream during Wet Weather using Shotgun Metagenomic Analyses. *Environ. Sci. Technol.* **52**, 9033–9044 (2018).

15. Senthil Kumar, P., Suganya, S. & Varjani, S. J. Evaluation of Next-Generation Sequencing Technologies for Environmental Monitoring in Wastewater Abatement. in 29–52 (Springer, Singapore, 2018). doi:10.1007/978-981-10-7485-1_3.

16. Chao, Y. *et al.* Metagenomic analysis reveals significant changes of microbial compositions and protective functions during drinking water treatment. *Sci. Reports 2013 31* **3**, 1–9 (2013).

17. Jia, S. *et al.* Metagenomic assembly provides a deep insight into the antibiotic resistome alteration induced by drinking water chlorination and its correlations with bacterial host changes. *J. Hazard. Mater.* **379**, (2019).

18. Garner, E. *et al.* Metagenomic Characterization of Antibiotic Resistance Genes in Full-Scale Reclaimed Water Distribution Systems and Corresponding Potable Systems. *Environ. Sci. Technol.* **52**, (2018).

19. Garner, E. *et al.* Impact of blending for direct potable reuse on premise plumbing microbial ecology and regrowth of opportunistic pathogens and antibiotic resistant bacteria. *Water Res.* **151**, 75–86 (2019).

20. Kim, S. *et al.* Transfer of antibiotic resistance plasmids in pure and activated sludge cultures in the presence of environmentally representative micro-contaminant concentrations. *Sci. Total Environ.* **468**–**469**, 813–820 (2014).

21. Huijbers, P. M. C., Flach, C. F. & Larsson, D. G. J. A conceptual framework for the environmental surveillance of antibiotics and antibiotic resistance. *Environ. Int.* **130**, (2019).

22. Manaia, C. M. *et al.* Antibiotic resistance in wastewater treatment plants: Tackling the black box. *Environ. Int.* **115**, 312–324 (2018).

23. Pruden, A., Bott, C., Blair, M. F., Miller, J. H. & Vaidya, R. Characterization of Organic Carbon and Microbial Communities for the Optimization of Biologically-Active Carbon (BAC) Filtration for Potable Reuse. *Water Res. Found. Proj. 4872/U1R16 Final Report.*

305-undefined (2020).

24. Sun, Y. *et al.* A pilot-scale investigation of disinfection by-product precursors and trace organic removal mechanisms in ozone-biologically activated carbon treatment for potable reuse. *Chemosphere* **210**, 539–549 (2018).

25. Vaidya, R. *et al.* Pilot Plant Performance Comparing Carbon-Based and Membrane-Based Potable Reuse Schemes. *Environ. Eng. Sci.* **36**, 1369–1378 (2019).

26. Vaidya, R., Wilson, C. A., Salazar-Benites, G., Pruden, A. & Bott, C. Implementing Ozone-BAC-GAC in potable reuse for removal of emerging contaminants. *AWWA Water Sci.* **2**, e1203 (2020).

27. Vaidya, R., Wilson, C. A., Salazar-Benites, G., Pruden, A. & Bott, C. Factors affecting removal of NDMA in an ozone-biofiltration process for water reuse. *Chemosphere* **264**, 128333 (2021).

28. Arango-Argoty, G. *et al.* MetaStorm: A Public Resource for Customizable Metagenomics Annotation. *PLoS One* **11**, (2016).

29. Jia, B. *et al.* CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573 (2017).

30. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).

31. Li, B. *et al.* Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *ISME J.* **9**, (2015).

32. Garner, E. *et al.* Metagenomic Characterization of Antibiotic Resistance Genes in Full-Scale Reclaimed Water Distribution Systems and Corresponding Potable Systems. *Environ. Sci. Technol.* **52**, 6113–6125 (2018).

33. R Core Team, 2018. R: A language and environment for statistical computing. *R Found. Stat. Comput. Vienna, Austria* (2018).

34. Oksanen, J. *et al.* Package 'vegan' Title Community Ecology Package Version 2.5-7. (2020).

35. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, 1–18 (2011).

36. Wickham, H. ggpolt2 Elegant Graphics for Data Analysis. *Use R! Ser.* 211 (2016).

37. Li, B. *et al.* Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *ISME J.* **9**, 2490–2502 (2015).

38. Luo, Y. *et al.* Trends in antibiotic resistance genes occurrence in the Haihe River, China. *Environ. Sci. Technol.* **44**, 7220–7225 (2010).

39. Pruden, A., Pei, R., Storteboom, H. & Carlson, K. H. Antibiotic Resistance Genes as

Emerging Contaminants: Studies in Northern Colorado†. *Environ. Sci. Technol.* **40**, 7445–7450 (2006).

40. Yin, X. *et al.* An assessment of resistome and mobilome in wastewater treatment plants through temporal and spatial metagenomic analysis. *Water Res.* **209**, 117885 (2022).

41. Sikorsky, J. A., Primerano, D. A., Fenger, T. W. & Denvir, J. DNA damage reduces Taq DNA polymerase fidelity and PCR amplification efficiency. *Biochem. Biophys. Res. Commun.* **355**, 431–437 (2007).

42. Li, W. & Sancar, A. Methodologies for Detecting Environmentally-Induced DNA Damage and Repair. *Environ. Mol. Mutagen.* **61**, 664 (2020).

43. Garner, E. *et al.* Metagenomic Characterization of Antibiotic Resistance Genes in Full-Scale Reclaimed Water Distribution Systems and Corresponding Potable Systems. *Environ. Sci. Technol.* **52**, 6113–6125 (2018).

44. Wan, K. *et al.* Accumulation of antibiotic resistance genes in full-scale drinking water biological activated carbon (BAC) filters during backwash cycles. *Water Res.* **190**, 116744 (2021).

45. Zhu, N. J., Ghosh, S., Edwards, M. A. & Pruden, A. Interplay of Biologically Active Carbon Filtration and Chlorine-Based Disinfection in Mitigating the Dissemination of Antibiotic Resistance Genes in Water Reuse Distribution Systems. *Environ. Sci. Technol.* **55**, 8329–8340 (2021).

46. Tan, Q. *et al.* Presence, dissemination and removal of antibiotic resistant bacteria and antibiotic resistance genes in urban drinking water system: A review. *Front. Environ. Sci. Eng.* **13**, 1–15 (2019).

47. Sha'arani, S. *et al.* Removal efficiency of Gram-positive and Gram-negative bacteria using a natural coagulant during coagulation, flocculation, and sedimentation processes. *Water Sci. Technol.* **80**, 1787–1795 (2019).

48. Hu, Y. *et al.* Occurrence and reduction of antibiotic resistance genes in conventional and advanced drinking water treatment processes. *Sci. Total Environ.* **669**, 777–784 (2019).

49. Krzemiński, P., Popowska, M., Krzemiński, P. & Popowska, M. Treatment Technologies for Removal of Antibiotics, Antibiotic Resistance Bacteria and Antibiotic-Resistant Genes. 415–434 (2020) doi:10.1007/978-3-030-40422-2_19.

50. Yang, L., Wen, Q., Chen, Z., Duan, R. & Yang, P. Impacts of advanced treatment processes on elimination of antibiotic resistance genes in a municipal wastewater treatment plant. *Front. Environ. Sci. Eng.* **13**, 1–10 (2019).

51. Xu, L., Campos, L. C., Canales, M. & Ciric, L. Drinking water biofiltration: Behaviour of antibiotic resistance genes and the association with bacterial community. *Water Res.* **182**, 115954 (2020).

52. Su, J. Q. *et al.* Antibiotic Resistome and Its Association with Bacterial Communities

during Sewage Sludge Composting. *Environ. Sci. Technol.* **49**, 7356–7363 (2015).

53. Yu, K. *et al.* Antibiotic resistome associated with microbial communities in an integrated wastewater reclamation system. *Water Res.* **173**, 115541 (2020).

54. Chen, Y. *et al.* High-throughput profiling of antibiotic resistance gene dynamic in a drinking water river-reservoir system. *Water Res.* **149**, 179–189 (2019).

55. Tong, J. *et al.* Microbial community evolution and fate of antibiotic resistance genes along six different full-scale municipal wastewater treatment processes. *Bioresour. Technol.* **272**, 489–500 (2019).

56. Sun, H. *et al.* Deciphering the antibiotic resistome and microbial community in municipal wastewater treatment plants at different elevations in eastern and western China. *Water Res.* **229**, 119461 (2023).

57. Jia, S., Bian, K., Shi, P., Ye, L. & Liu, C. H. Metagenomic profiling of antibiotic resistance genes and their associations with bacterial community during multiple disinfection regimes in a full-scale drinking water treatment plant. *Water Res.* **176**, 115721 (2020).

58. Czekalski, N. *et al.* Inactivation of Antibiotic Resistant Bacteria and Resistance Genes by Ozone: From Laboratory Experiments to Full-Scale Wastewater Treatment. *Environ. Sci. Technol.* **50**, 11862–11871 (2016).

59. Alexander, J., Knopp, G., Dötsch, A., Wieland, A. & Schwartz, T. Ozone treatment of conditioned wastewater selects antibiotic resistance genes, opportunistic bacteria, and induce strong population shifts. *Sci. Total Environ.* **559**, 103–112 (2016).

60. Zheng, J. *et al.* Effects and mechanisms of ultraviolet, chlorination, and ozone disinfection on antibiotic resistance genes in secondary effluents of municipal wastewater treatment plants. *Chem. Eng. J.* **317**, 309–316 (2017).

61. Stapf, M., Miehe, U. & Jekel, M. Application of online UV absorption measurements for ozone process control in secondary effluent with variable nitrite concentration. *Water Res.* **104**, 111–118 (2016).

62. Zanacic, E., Stavrinides, J. & McMartin, D. W. Field-analysis of potable water quality and ozone efficiency in ozone-assisted biological filtration systems for surface water treatment. *Water Res.* **104**, 397–407 (2016).

63. Li, L. G., Huang, Q., Yin, X. & Zhang, T. Source tracking of antibiotic resistance genes in the environment — Challenges, progress, and prospects. *Water Res.* **185**, 116127 (2020).

64. Xiang, Q. *et al.* Seasonal change is a major driver of soil resistomes at a watershed scale. *ISME Commun. 2021 11* **1**, 1–7 (2021).

65. Wan, K., Lin, W., Zhu, S., Zhang, S. & Yu, X. Biofiltration and disinfection codetermine the bacterial antibiotic resistome in drinking water: A review and meta-analysis. *Front. Environ. Sci. Eng. 2019 141* **14**, 1–12 (2019).

66. Ju, F. *et al.* Wastewater treatment plant resistomes are shaped by bacterial composition, genetic exchange, and upregulated expression in the effluent microbiomes. *ISME J. 2018 132* **13**, 346–360 (2018).

67. Hu, Q. *et al.* Metagenomic insights into ultraviolet disinfection effects on antibiotic resistome in biologically treated wastewater. *Water Res.* **101**, 309–317 (2016).

68. Lira, F., Vaz-Moreira, I., Tamames, J., Manaia, C. M. & Martínez, J. L. Metagenomic analysis of an urban resistome before and after wastewater treatment. *Sci. Reports 2020 101* **10**, 1–9 (2020).

CHAPTER 4: ADVANCING METAGENOMICS AS A TOOL FOR PROFILING MICROBIAL FUNCTIONAL CAPACITY: INSIGHT INTO AN OZONE- BIOLOGICALLY ACTIVE CARBON FILTRATION POTABLE REUSE TREATMENT TRAIN

ABSTRACT

Metagenomic sequencing is a promising tool for revealing mechanistic insight into biological treatment. Such insights are especially needed to support advanced potable water reuse treatments that employ biological processes as a means to maximize removal of recalcitrant trace organic carbon constituents, such as 1,4-dioxane. Profiling functional genes captured by metagenomic sequencing could also provide fundamental insights into other treatment processes, such as disinfection. Here, we comprehensively profiled functional genes of interest through each stage of a carbon-based potable reuse train and contextualized shifts in spatial and temporal variation to changes in operational conditions and water quality parameters. This approach made it possible to relate the composition and abundances of functional genes present to individual treatment processes. Consistent with expectations, gene abundances associated with metabolic processes were highest during biologically-active carbon filtration. However, the fact that differentially abundant genes associated with metabolic processes were greatest during GAC contacting suggests that this treatment stage provides biological treatment, in addition to removal via sorption. Overall, profiling of functional genes via metagenomics proved valuable in assessing the biological treatment capacity of various treatment stages and served to identify numerous genes associated with specific functions of interest, including biofilm formation, nitrogen cycling, metabolic processes, DNA repair, monooxygenase activity, etc.

INTRODUCTION

As water and wastewater sectors continue to shift toward incorporating innovative and sustainable water reuse treatment processes, there is a need to keep pace with a mechanistic understanding of how these processes function and their overall treatment capacities. In particular, biologically active filtration (BAF) is increasingly being viewed as a viable alternative to reverse osmosis based water reuse treatment trains. In addition to lack of production of a brine waste stream, an advantage of BAF over reverse osmosis is that organic contaminants can be further broken down by microbial communities. Typically, BAF is preceded by an advanced oxidative processes (AOPs) to improve the biodegradability of recalcitrant compounds. AOPs can also contribute to pathogenic removal. Given that microbial communities are the fundamental driver of treatment achieved by AOP-BAF systems, there is a need to better characterize their capacity for biodegradation and pathogen attenuation.

Previous characterizations of BAF performance have largely been achieved through monitoring reductions in contaminant and nutrient concentrations or bulk, low resolution assessments of total microbial activity or metabolism[1–6]. However, little work has been done to characterize or establish a fundamental understanding of genes encoding key metabolic functions of interest, such as trace organic carbon (TrOC) or nutrient breakdown, within BAF units. This is especially important when beginning to assess carbon-based, advanced treatment trains for the removal of TrOCs that pose known or potential human health concerns; including industrial contaminants, pharmaceuticals, personal care products, and other contaminants of emerging

concern (CECs), that may be subject to current and future water reuse regulations. Beyond TrOC reduction, microbial communities active within BAF treatments involved in stabilizing nitrogen in the finished water are also of interest. This includes canonical nitrification pathways, comammox (i.e., complete nitrification by converting ammonia into nitrite and then nitrate), or anammox (organisms that can anaerobically covert nitrite and ammonium into nitrogen gas and water), or even yet unidentified organisms that may carry out relevant metabolic processes. Additionally, heterotrophic bacteria can also impart some degree of nitrogen and phosphorus assimilation via anabolic pathways. However, the precise role of BAF microbial communities in organic contaminant and nutrient removal is not well understood. In particular, complex microbial dynamics -and their impact on treatment performance- are difficult to characterize and can potentially be influenced by upstream stress conditions, chemical additions [7], operational conditions [8], supplemental nutrient sources [9], and general changes in influent substrate makeup and overall water quality.

Next generation sequencing (NGS), especially shotgun metagenomic sequencing, could potentially serve as a powerful tool for profiling the functional capacity of biologically complex processes as it relates to upstream conditions and overall treatment performance. NGS has, to some extent, previously been applied to study functionally related activities and metabolic processes in drinking water and wastewater environments. Most of these studies have focused on profiling general functional capacity or target specific functions of interest, such as: nitrification and denitrification processes [10,11], catabolic biodegradation of compounds [12,13], methanogenesis [14,15], the impact of disinfection [16,17], and biofilm interaction [18,19], among others. However, a significant knowledge gap still remains for potable reuse applications, especially BAF-related processes. Thus far, metagenomic profiling of functional genes has mainly been applied to provide complementary information about relatively small subsets of samples with limited replication. Prior studies have mainly emphasized singular treatment processes, most notably activated sludge reactors and anaerobic digesters. The lack of mechanistic understanding of multiple sequential treatment processes is especially notable in the case of advanced water treatment. The implementation of metagenomic profiling of functional genes to better understand coupling of AOPs with carbon-based biological filtration is notably lacking, especially as the prospects for more sustainable reuse options continue to grow within the water and wastewater industry. There is also a need to for more integrated analysis of water quality, operational condition, and metagenomic functional data to provide a more holistic understanding of water and wastewater treatment systems.

Here we utilized shotgun metagenomics sequencing to profile and track specific categories of functional genes of interest throughout a carbon-based potable reuse train employing coagulation, flocculation, and sedimentation (Floc-Sed), ozonation, biologically-active carbon (BAC) filtration, granular active carbon (GAC) contacting, and ultraviolet (UV) disinfection. Specifically, we tracked genes related to biofilm formation that might impact the effectiveness of settling processes (e.g. coagulation, flocculation, and sedimentation), oxidative process (e.g. ozonation), or disinfection processes (e.g. UV disinfection). Oxidative and disinfection process are also concerned with the presence of stress response and repair mechanisms available to microorganisms. Further, we elevated genes related to numerous metabolic processes to better understand degradation. Shifts in the spatial and temporal variation associated with these functional genes were related to changes in operational conditions and water quality parameters. The overarching objective of this was to develop and a metagenomic-

based functional gene profiling approach and demonstrate its application towards holistic analysis of the functional capacity of an AOP/BAF potable reuse train. Specifically, this works objectives were to: 1) characterize dynamic shifts in functional profiles both spatially, between treatments, and temporally, over continued operation, 2) identify, quantify, and compare functional genes of interests related to important CEC degradative pathways, observed bulk TrOC and specific CEC reductions, and corresponding operational conditions, 3) characterize functional genes of interest related to important nutrient cycling processes, 4) identify the impact of disinfection process and changes in operation on stress response genes, and repair mechanisms, and 5) utilize functional annotation to assess additional functions of interest, including: horizontal gene transfer, and biofilm development.

METHODS

Site Description, Sample Collection, and Preservation

Figure 4.1 provides a schematic of a 4.3 gpm pilot-scale facility associated with the SWIFT initiative. Bulk water samples from each stage of pilot treatment were collected over an ~18-month period. This pilot treated tertiary (denitrified secondary) effluent and employs: Floc-Sed, ozonation, BAC filtration, GAC contacting, and UV disinfection. Both BAC filters and GAC contactors were operated with parallel process at shorter (BAC5, GAC10) and longer (BAC10, GAC20) EBCT. Various chemicals and operational conditions were also adjusted periodically with major changes and upset conditions outlined in SI Table A.1. Specifically, individual treatment processes were exposed to suppressed temperatures, variable ozone dosing, multiple forms of chlorine quenching, addition of filter aid polymer, and dosing of supplemental nutrients prior to BAC treatment. Additionally, a few upset conditions were also identified, including loss of BAC media and elevated turbidity following Floc-Sed treatment. Additional details on SWIFT pilot operation during this study period are available in prior publications [20–24] and the SI.

On site sampling consisted of triplicate bulk water samples collected in sterilized 1-L polypropylene bottles. Collected samples were then transported back to Virginia Tech on ice and were processed within 24 h of collection. Processing consisted of filter concentration onto 0.22-µm mixed cellulose esters membrane filter (Millipore, Billerica, MA), fragmentation via sterilized tweezers, and storage at – 20°C. Filtering was completed until clogging with exact sample volumes varying with the total suspended solid content of the different sampling locations. Exact filter volumes were recorded at the time of processing and typically ranged between 0.2 (Pilot influent) to 1 liter (all other sampling locations). DNA was extracted from filters using a FastDNA SPIN Kit for Soil (MP Biomedicals, Solon, OH) according to the manufacturer's instructions with extracts being stored at -20 °C prior to downstream molecular analysis.

Shotgun Metagenomic Sequencing and Analysis

A subset of 72 samples, collected over 10 sampling events, were selected and submitted for shotgun metagenomic sequencing via three individual lanes. Quantification of extracted DNA was completed using the Qubit dsDNA HS-assay (Invitrogen™ Qubit™ 3 Fluorometer, Waltham, MA) and samples were removed from consideration if their concentrations were below 0.2 ug/L. Prior to sequencing, triplicate samples had their raw DNA pooled together, at equal

mass, to reach their target concentration. Sequencing was conducted on an Illumina NextSeq500 (Illumina, San Diego, CA, United States) using 2x75 paired end reads and the NEB Ultra II library prep by The Scripps Research Institute (The Scripps Research Institute, San Diego, CA, United States). The NEB Ultra II library prep was selected due to its ability to handle samples with DNA masses down to 1ng and the low concentration of DNA associated with ozone and UV effluents present in the study's sampling matrix.

Following sequencing, reads were uploaded, read matched, and annotated using the Metastorm platform [25] and default parameters. The target depth corresponded to approximately 17 million reads per samples, with additional metagenomic read statistics detailed in SI Figure C.1 through C.4. The present analysis utilizes an uploaded version of the UniProtKB-Swiss-Prot (version from June 2020) reviewed database [26] for annotation. Functional genes relative abundances were normalized with 16S rRNA gene counts using the Silva 16S rRNA database [27] (i.e., ARG copies per copies of 16S rRNA genes generated from metagenomic data) [28].

Data and Statistical Analysis

Functional gene annotations were imported and analyzed in R [29] using RStudio (RStudio Team, 2019). Annotated genes were assessed for relevance to the study objectives and were subsequently categorized into groups of interest, filtering out irrelevant annotations. To achieve this, annotated gene-ids from UniProtKB-Swiss-Prot were accompanied by Gene Ontology terms (GO) from the GO knowledge base [30,31]. These functional classifications were utilized to categorize genes into groups of similar functions, filter for functions of interest, and assess changes in functional capacity of collected samples. The R package GO.db [32] and relationships presented at the EMBL European Bioinformatic Institute (EMBL-EBI) was utilized to identify GO-terms of interest and all children and offspring associated with these functional targets. All gene-ids only associated with Eukaryotes were removed from the dataset.

The GO-term groups were used to further classify the initial annotations from UniProtKB-Swiss-Prot and compare the relative gene abundances as they related to spatial and temporal shifts. Specifically, all relevant functional genes associated with high level biological processes were classified into the respective child terms associated with biological processes (GO:0008150, e.g. viral process, response to stimulus, metabolic process, etc.) and metabolic processes (GO:0008152, e.g. biosynthetic process, nitrogen compound metabolic process, antibiotic metabolic process, etc.). Additionally, higher resolution GO-terms, and their offspring, associated with specific functions of interest were also assessed in a similar manner. These included: horizontal gene transfer, biofilm formation, DNA repair, response to ozonation, monooxygenase activity, and nitrogen cycling. After sub-setting the entire functional annotation for gene categories of interest, data analysis and statistical analysis was conducted in association with stated research objectives.

Functional profiles of interest were assessed using Bray-Curtis dissimilarities generated in R to compare changes in beta-diversity between individual treatment processes, operational conditions, and temporal variability. Statistical testing associated with beta-diversity was conducting utilizing the analysis of similarities (ANOSIM) for categorical variables and ADONIS test for continuous variables, both using the vegan package [33]. Groups of categorized functional genes were assessed using paired Wilcoxon rank-sum tests in R with Benjamini-

Hochberg p-value correction to identify differences in abundances between tested conditions and treatment processes.

Discriminatory abundant functional genes were identified using the DESeq2 package [34] in R using adjusted p-value cutoffs of 0.05. Discriminatorily abundant genes were calculated between (1) each treatment process vs all other treatment processes and (2) between the influent and effluent of each individual treatment processes. Discriminatory functional genes from (2) were also compared utilizing upset graphs generated in the complexUpset package [35].

Correlation analysis was conducting using the r-corr function in the Hmisc package [36] using Benjamini-Hochberg correction and p-value cutoffs of 0.10 for more exploratory correlations. Normality was checked using the Shapiro-Wilk's test. Correlation analysis with normally distributed variables were tested with Pearson tests and non-parametric testing was conducted via Spearman. Statistical significance was set at $p < 0.05$, unless otherwise noted, with full statistical reporting presented in SI spreadsheet. All figures were produced utilizing the package ggplot2 [37] in R, unless noted otherwise.

RESULTS

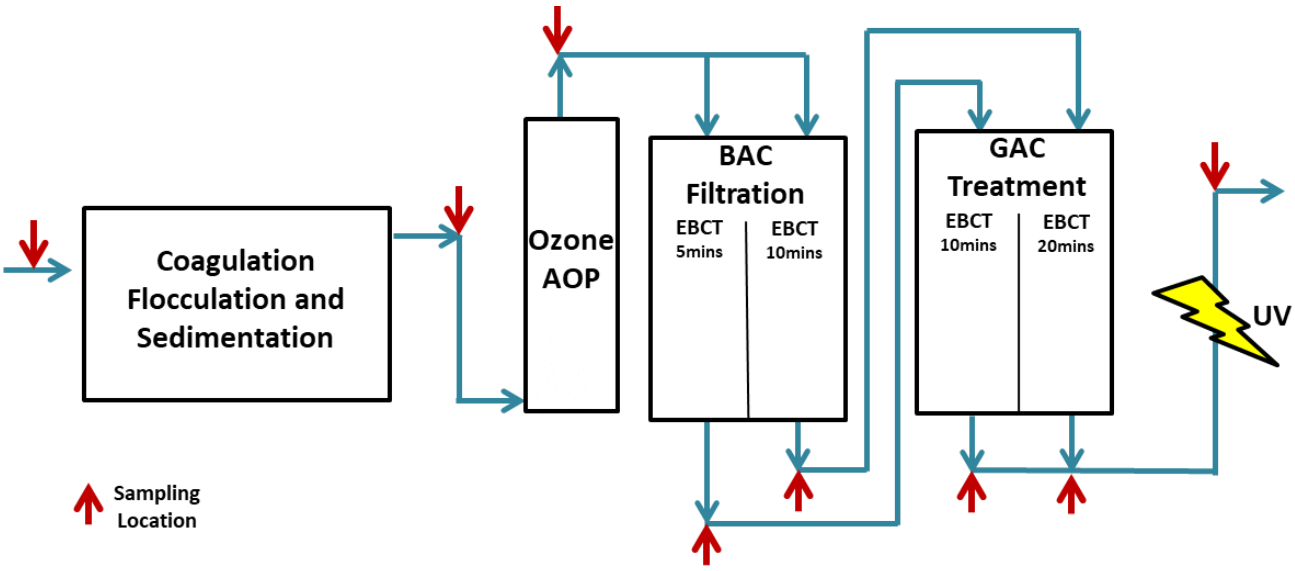


Figure 4.1: HRSD SWIFT Pilot Facility's process schematic. Red arrows indicate sampling locations. The pilot influent consisted of effluent from the York River wastewater treatment plant at a flowrate of 4.3 gallons per minute.

Assessment of Comprehensive Functional Profiles

All functional genes annotated using the UniProt database (except those associated as Eukaryotes) were used to assess the functional profile within and throughout each stage of advanced treatment. NMDS plots generated via Bray Curtis dissimilarity metrics graphically represent shifts in these functional profiles. Figure 4.2a provides the relationships of these functional profiles at each stage of treatment using gene counts normalized to 16S rRNA gene abundances determined via shotgun metagenomic sequencing, while Figure 4.2b presents similar relationships during only BAC filtration and GAC contacting. A similar analysis was conducted

including only genes associated with any metabolic functions (determined by GO classifications), but failed to provide appreciable differences at a similarly high-level classification (SI Figure E.1).

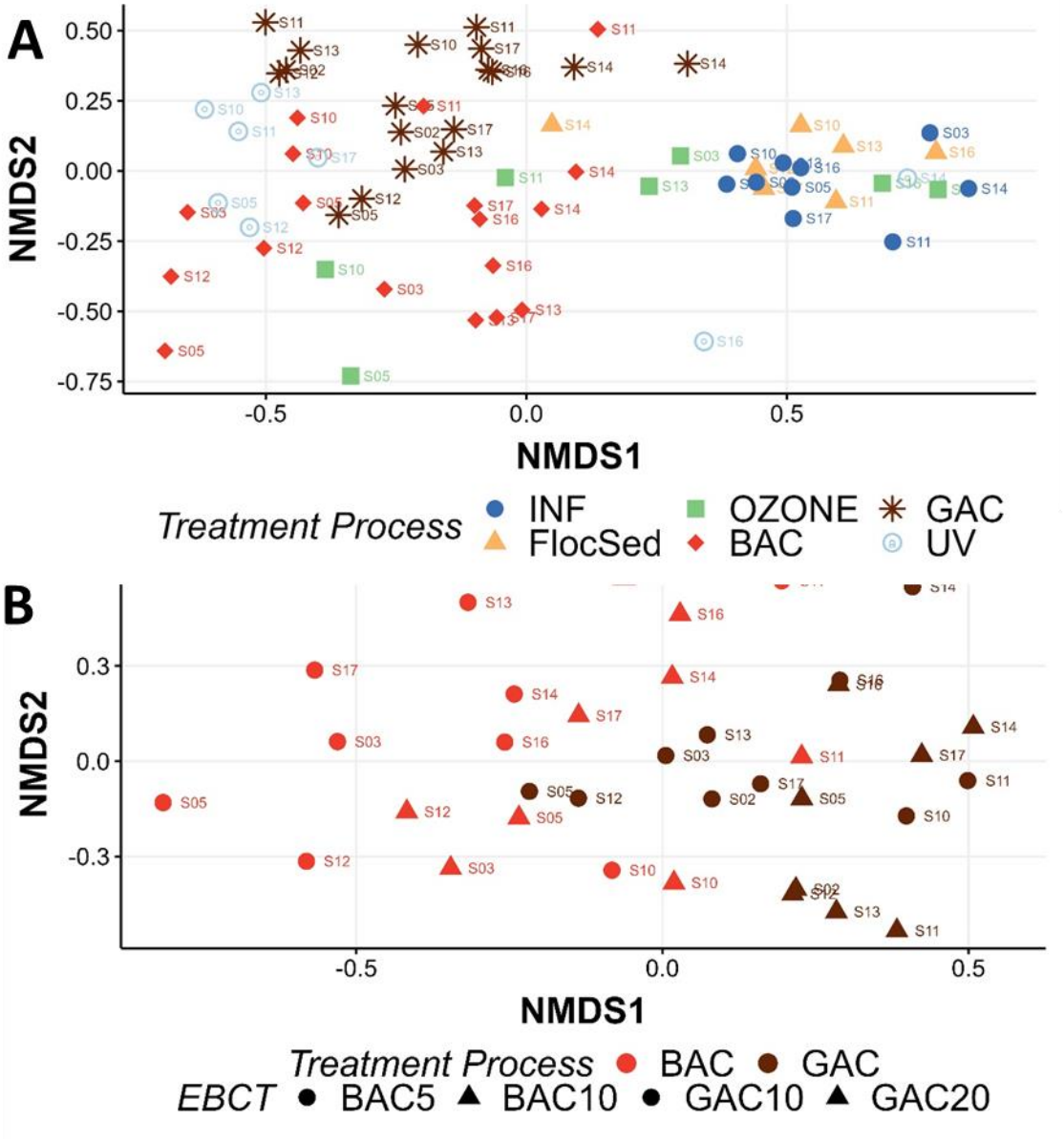


Figure 4.2: NMDS analysis of all annotated functional genes not associated with Eukaryotic organisms, based on gene counts normalized to 16S rRNA gene copies, for (A) all treatment processes (ANOSIM, p-value<0.05, r-stat = 0.59), and (B) BAC filtration and GAC contracting at tested EBCTs (ANOSIM, p-value<0.05, r-stat = 0.33).

From Figure 4.2a, treatment processes clearly shaped the corresponding functional gene profile, both with (ANOSIM, p-value<0.05, r-stat = 0.50) and without (ANOSIM, p-value<0.05, r-stat = 0.59) accounting for differences in EBCT, moreso than any other tested factor. Further, statistically significant shifts between the functional profiles were identified between

influent:ozone (ANOSIM, p-value<0.05, r-stat =0.34), ozone:BAC5 (ANOSIM, p-value<0.05, r-stat =0.25), ozone:BAC10 (ANOSIM, p-value<0.05, r-stat =0.33), ozone:BAC (ANOSIM, p-value<0.05, r-stat =0.37), BAC5:GAC10 (ANOSIM, p-value<0.05, r-stat =0.35), BAC10:GAC20 (ANOSIM, p-value<0.05, r-stat =0.36), BAC:GAC (ANOSIM, p-value<0.05, r-stat =0.33), GAC:UV (ANOSIM, p-value<0.05, r-stat =0.47), and influent:UV (ANOSIM, p-value<0.05, r-stat =0.62). Statistically significant differences between the functional profiles were not identified between influent:FlocSed, FlocSed:ozone, BAC5:BAC10, and GAC10:GAC20.

Variability of the functional profiles within treatment processes were most associated with addition of supplemental nutrients, temperature, pilot runtime, and nitrite adjusted O3:TOC ratios, while other operational condition changes (e.g. addition of filter aid polymer, and chlorine quenching type) failed to result in significant changes to the full functional profile during the course of this experiment. Specifically, the addition of supplemental nutrients resulted in statistically significant different functional profiles within the BAC filters (ANOSIM, p-value<0.05, r-stat =0.25), but not the GAC contactors. Nitrite adjusted O3:TOC ratios were also found to correspond well to functional variability associated with ozonation (Adonis, p-value<0.05, R2 =0.31), BAC filtration (Adonis, p-value<0.05, R2 =0.12), GAC contactors (Adonis, p-value<0.05, R2 =0.11), and UV disinfection (Adonis, p-value<0.05, R2 =0.28). Temperature was well related with the combined process prior (Adonis, p-value<0.05, R2 =0.12) and BAC filters (Adonis, p-value<0.05, R2 =0.12). Both BAC filters (Adonis, p-value<0.05, R2 =0.72) and GAC contactors (Adonis, p-value<0.05, R2 =0.61) were also related to pilot runtime.

Analysis of Functional Genes Classified into Biological and Metabolic Processes

Annotated functional genes were classified into the highest-level child terms associated with GO-terms focused on biological (GO:0008150) and metabolic (GO:0008152) processes. SI Figure E.2 provides all child categories associated with biological processes while Figure 4.3 provides selected classifications of interest. Initial analysis at this high level of resolution indicated few distinctions between BAC filters and GAC contactors operated at different EBCT, therefore analysis presented does not separate them.

From Figure 4.3, no statistically significant differences between influent and FlocSed were found among the selected categories. Biological regulation (GO:0065007), and reproduction (GO:0000003) were found to be enriched between influent and ozonation (Wilcoxon, p-value < 0.05), while enrichment between influent and BAC filtration was starker for biological regulation, metabolic processes (GO:0008152), reproduction, and response to stimulus (GO:0050896) (Wilcoxon, p-value < 0.05). Similar statistically significant differences were observed for biological regulation, metabolic processes, reproduction, and response to stimulus classifications between both BAC and GAC treatments and influent and UV effluent (Wilcoxon, p-value < 0.05). BAC and GAC treatments also indicated differences within the detoxification (GO:0098754) classification (Wilcoxon, p-value < 0.05). Statistically significant differences were also measured between GAC contacting and UV treatment for the biological regulation, detoxification, metabolic processes, and response to stress classified functional genes. Interestingly, genes associated with viral processes (GO:0016032) were found to be highest at the inlet of the pilot train and experienced reduction during ozonation and BAC treatment before continuing at lower concentrations during GAC contacting and UV disinfection. This shift resulted in a statistically significant decrease between the influent and BAC effluent (Wilcoxon,

p-value < 0.05). Overall, metabolically related functional genes made up the vast majority of 16S rRNA normalized gene counts, followed by biological regulation and stimulus response.

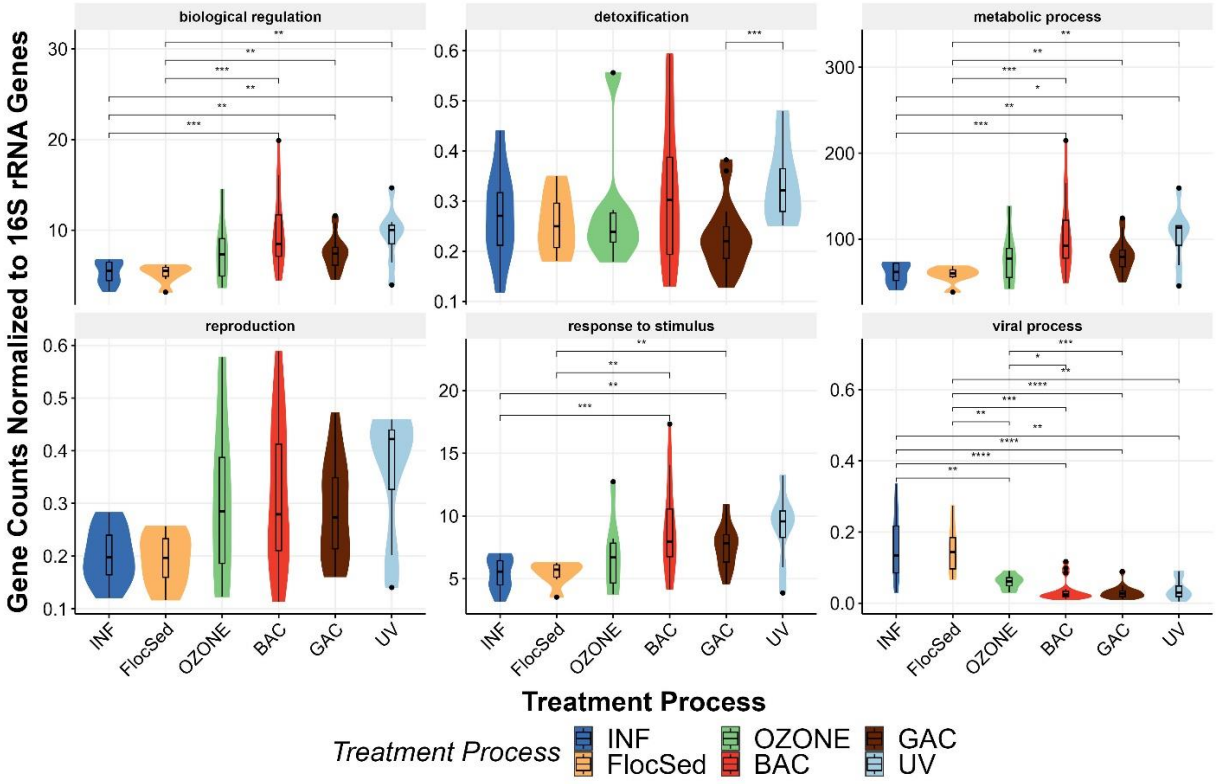


Figure 4.3: Violin plots of functional genes counts normalized by 16S rRNA genes grouped by treatment process and selected categories of GO-child terms related to biological process (GO:0008150)

SI Figure E.3 provides all child categories associated with metabolic processes while Figure 4.4 provides selected classifications of interest. Overall, most high-level categories associated with metabolic processes displayed similar relationships between treatment processes. Specifically, significant differences were observed in biosynthetic (GO:0009058), catabolic (GO:0009056), cellular metabolic (GO:0044237), nitrogen compound metabolic (GO:0006807), organic substance metabolic (GO:0071704), primary metabolic (GO:0044238) and small molecule (GO:0044281) metabolic processes (Wilcoxon, p-value < 0.05) between the influent:BAC, BAC:GAC, GAC:UV, and influent:UV. The same relationships were observed among secondary metabolic processes, in addition to differences between ozonation and BAC filtration (Wilcoxon, p-value < 0.05). Antibiotic metabolic process (GO:0016999), on the other hand, only experienced statistically difference relationships between process prior to ozonation and BAC filtration and influent and UV effluent (Wilcoxon, p-value < 0.05). Similarly to the full characterization of biological processes, no significant difference between influent and FlocSed effluent nor an impact of EBCT on either BAC filtration or GAC contacting were found for genes associated with metabolic process. Although there appeared to be a different trend between ozonation and BAC processes, with BAC filtration yielding increased abundances, it was not statistically significant.

Of all classified categories, normalized gene abundances were greatest within biosynthetic, cellular metabolic, nitrogen compound metabolic, organic substance, primary metabolic and small molecule metabolic processes. Conversely, fewer normalized gene abundances were associated with antibiotic metabolic, secondary metabolic, and catabolic processes.

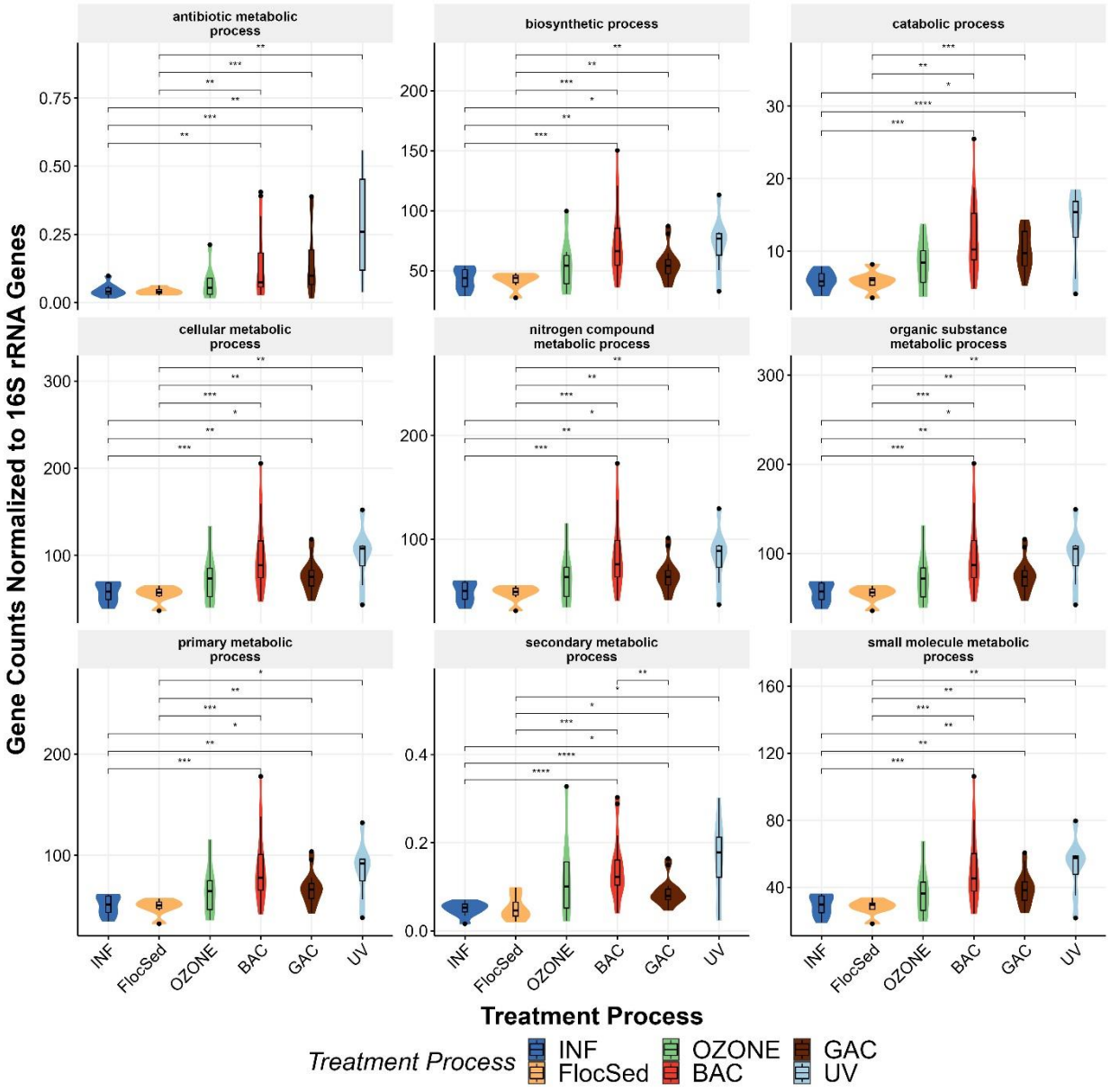


Figure 4.4: Violin plots of functional genes counts normalized by 16S rRNA genes grouped by treatment process and selected categories of GO-child terms related to metabolic process (GO:0008152)

Identification of Differentially Abundant Functional Genes between Treatment Processes

SI Table E.2 provides summary counts of statistically significant, unique, and differentially abundant functional genes (DeSeq2, p-value<0.05) when comparing specified treatment types to all other samples. From SI Table A.1, the FlocSed (35 unique differentially abundant genes) and Ozonation (34) processes had the lowest number of associated, differentially abundant genes. FlocSed also had the lowest percentage of differentially abundant genes (2.5%) associated with that treatment process compared to all other samples, while ozonation experienced the second highest percentage (64.2%). GAC contacting (11698) had, by far, the highest number of unique, differentially abundant functional genes as well as the highest percentage of genes associated with that treatment process (69.3%) followed by BAC filtration (4239, 33.0%), influent (2623, 24.2%) and UV disinfection (464, 29.8%). When combined, influent and FlocSed processes (sometimes referred to as 'Front') resulted in increased quantities and proportions of location specific, differentially abundant genes (5638, 34.1%). Combinations of ozonation and UV disinfection resulted in higher quantities of associated genes with a lower proportion (794, 25%) while combination of BAC and GAC treatments (sometimes referred to as BAF) resulted in similar numbers to that of GAC alone (11614, 62.1%).

When categorized by biological processes (SI Figure A.1 and Figure 4.5) counts of unique, differentially abundant genes were found to be much higher during BAC and GAC treatment for biological regulation, detoxification, metabolic processes, reproduction, and response to stimulation, with the greater counts experienced during GAC contacting. These categories also typically found slightly fewer genes associated with influent than BAC filtration with the combination of influent and FlocSed experiencing slightly more than BAC filtration. Conversely, unique, differentially abundant genes associated with viral processes were much greater in at the front of the treatment train than BAC or GAC effluents. In all cases, ozonation and UV disinfection account for very few categorized genes, if any.

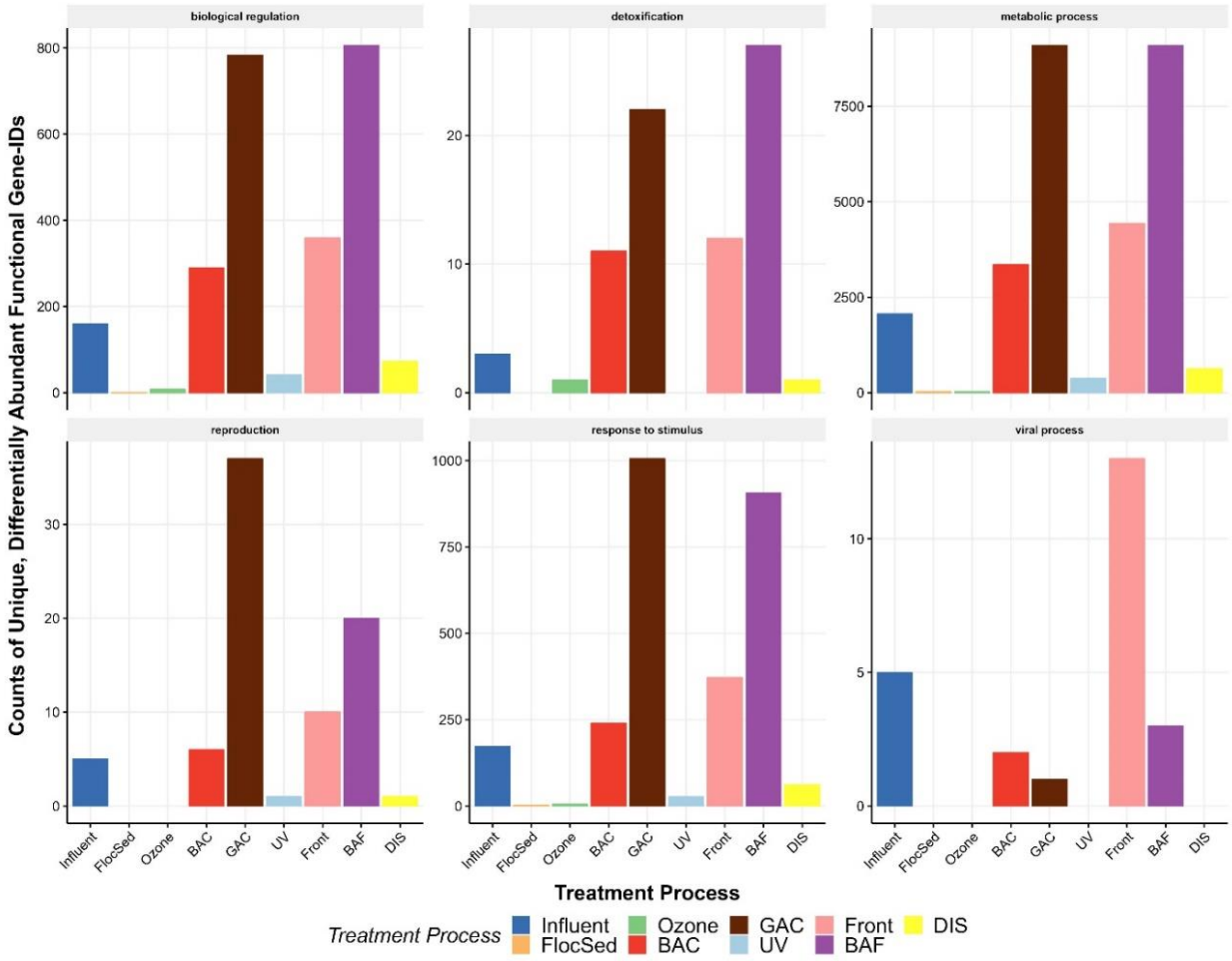


Figure 4.5: Counts of unique, differentially abundant genes determined by DESeq2 with an adjusted p-value cutoff of 0.05. Gene counts are grouped by treatment process and selected categories of GO-child terms related to biological process (GO:0008150). Differential genes were determined by comparisons between individual treatment process and all other samples.

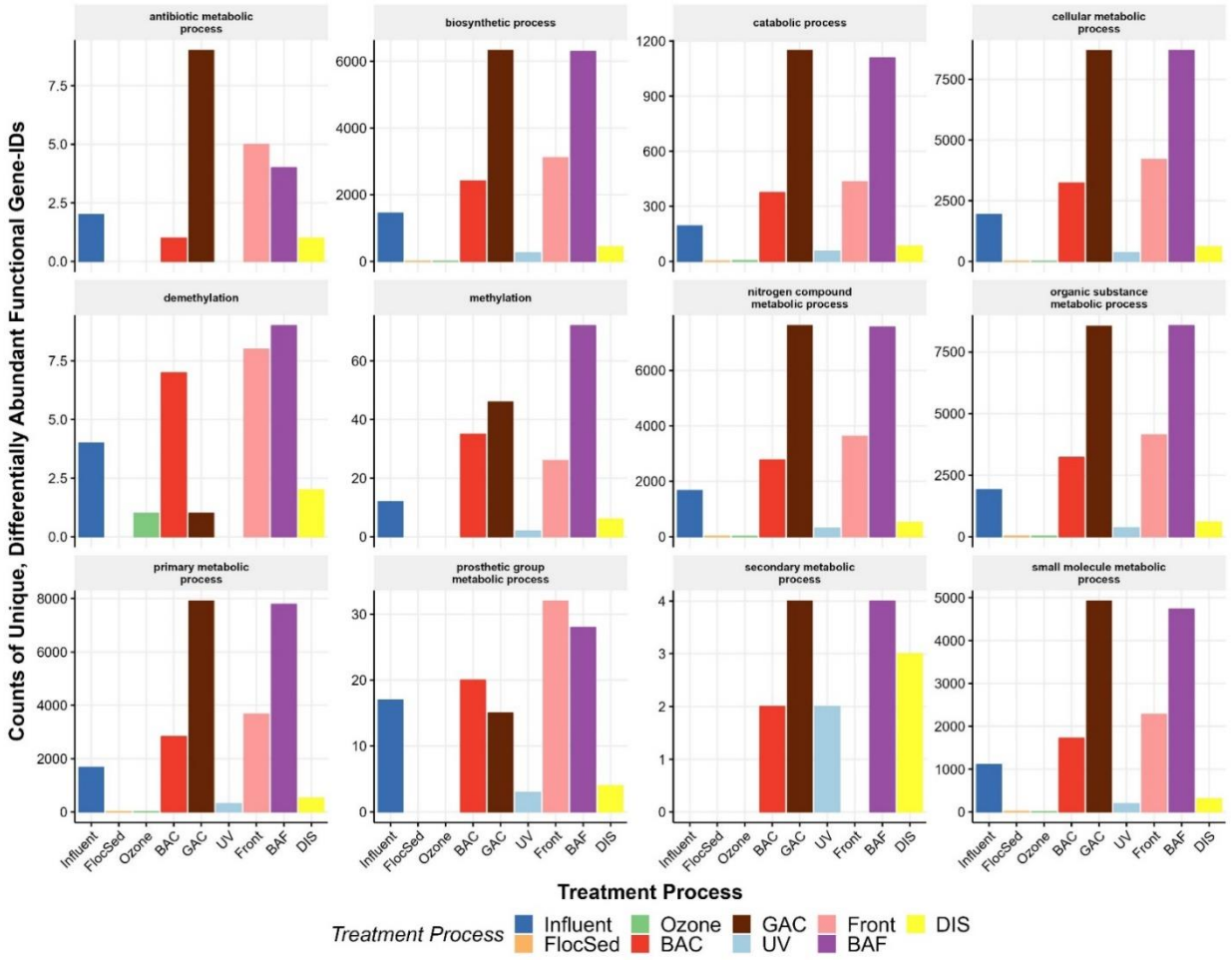


Figure 4.6: Counts of unique, differentially abundant genes determined by DESeq2 with an adjusted p-value cutoff of 0.05. Gene counts are grouped by treatment process and selected categories of GO-child terms related to metabolic process (GO:0008152). Differential genes were determined by comparisons between individual treatment process and all other samples.

When categorized by metabolic processes (SI Figure A.2 and Figure 4.6) counts of unique, differentially abundant genes followed similar trends as the biological processes' categorization. Notably, GAC contacting contained the highest number of unique genes per category followed by BAC filtration and influent. The combination of front processes was also found to be slightly higher than BAC. These categories include biosynthetic, catabolic, cellular metabolism, nitrogen compound metabolic, organic substance metabolic, primary metabolic, and small molecular metabolic processes. Of these, primary metabolic, organic substance metabolic, cellular metabolic, biosynthetic, nitrogen compound metabolic, and small molecular metabolic processes experienced the highest number of unique, discriminatorily abundant functional genes. Unique genes associated with methylation were found in higher abundance than demethylation with the later experiencing a greater number of genes within the BAC filters and treatment processes at the front of the pilot. Conversely, the higher number of unique genes associated with methylation experienced lowest abundances early in the treatment train and peaked during GAC contacting and BAF (combination of BAC and GAC samples). Unique functional genes

associated with prosthetic group metabolic processes were found at highest abundances within the BAC and influent before decreasing during GAC contacting. Similarly, grouped comparisons found prosthetic group metabolic processes to contain more unique functional genes prior to ozone than during BAF. Secondary metabolic processes made up a small fraction of the observed, unique functional genes with no differential genes identified prior to BAC filtration. Instead, the number of unique functional genes were highest during GAC treatment and the same during BAC filtration and UV disinfection. The total number of unique, discriminatorily abundant functional genes mirrored that of the trends seen with mean abundances for most categories (SI Figure A.3 and SI Figure A.4). However, this was not the case with antibiotic metabolic processes nor secondary metabolic processes. Antibiotic metabolic processes experienced more discriminatorily, mean gene abundances during BAF than another other treatment process or combination of processes. Meanwhile, secondary metabolic processes experienced similarly high mean abundances of unique genes associated with BAC treatment, with higher abundances during UV disinfection than GAC contacting.

Though this analysis was effective at identifying unique, discriminatorily abundance functional genes at each stage of treatment, it failed to identify the fate of discriminatorily abundant genes as they passed through multiple stages of treatment. For this purpose, statistically significant, discriminatorily abundant functional genes were identified using the influent and effluent of each respective treatment train. Figure 4.7 provides the relationships of those statistically significant, unique, discriminatorily abundant functional genes in relation to unique, discriminatorily abundant functional genes from other treatment processes. From Figure 4.7, UV effluent had the greatest number of non-shared differentially abundance functional genes and the second highest total number of genes between all treatment processes. GAC10 was identified as having the highest number of total genes and second highest number of non-shared genes. BAC10 and GAC20 had the next highest number of non-shared and total unique genes with both values well below that of UV effluent and GAC10. FlocSed effluent was found to have the fewest number of total, non-shared, and shared genes. BAC5 and ozone effluent had similar levels of non-shared genes while ozone shared very few genes with any treatment process other than UV disinfection while BAC5 shared a very high proportion of its genes with other treatment processes, especially BAC10. BAC and GAC treatments operated at different EBCT were identified as having the greatest number of shared genes between any combination of treatment processes. Interestingly, BAC10 also shared a high number of genes with the GAC contactors operated at multiple EBCT, specifically: GAC20 (2), GAC10 (86), GAC10 and GAC20 (17), and all other BAF process (1). This was much higher than relationships of BAC5 with the GAC contactors where only a single discriminatory gene was shared between any combination.

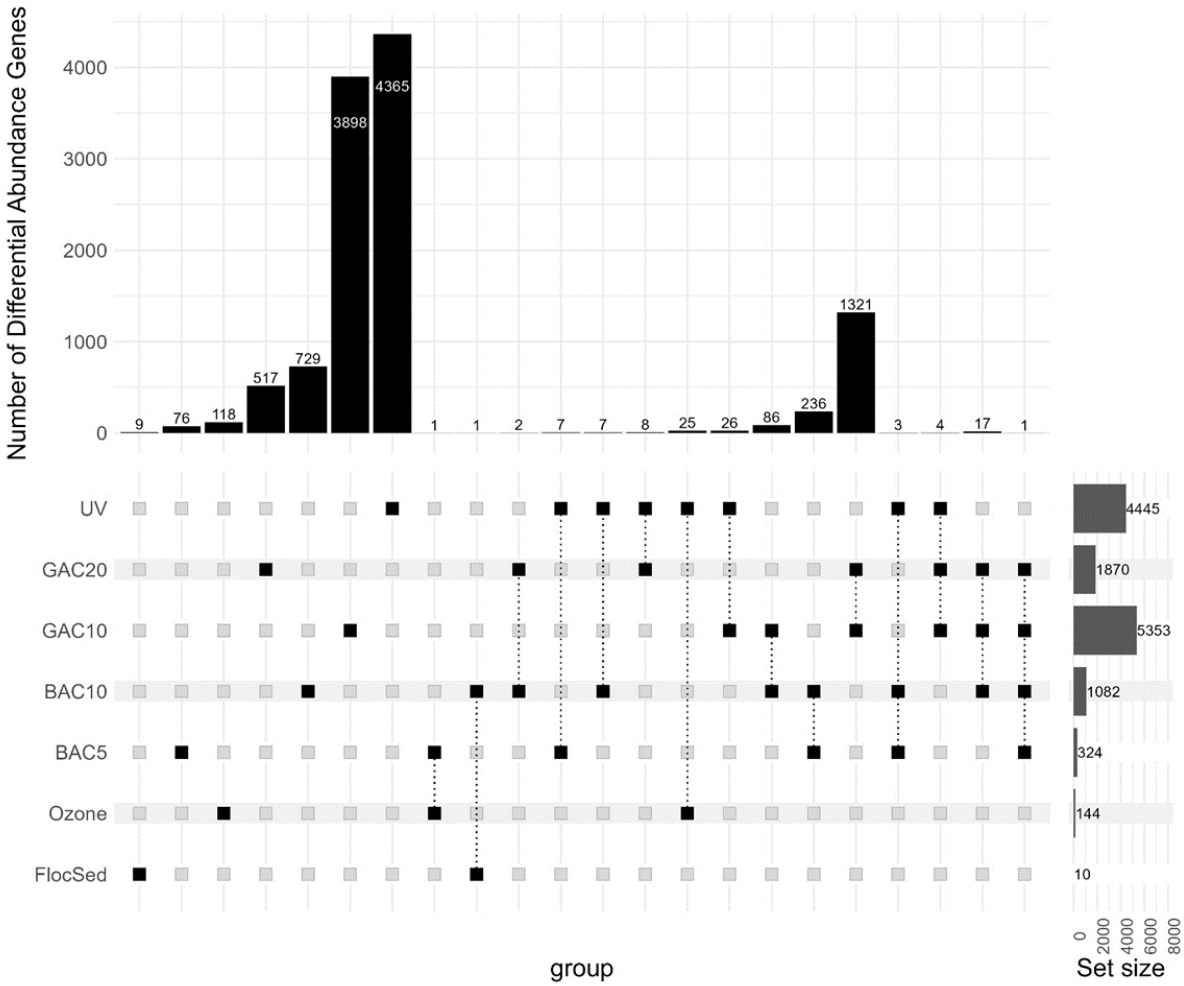


Figure 4.7: Provides a breakdown of all unique, differentially abundant genes (determined by DESeq2 with an adjusted p-value cutoff of 0.05) identified throughout treatment and the relationships of those differentially abundant functional genes to other treatment processes. Differential genes were determined by comparisons between the influent and effluent of each treatment process.

Analysis of Targeted Functions of Interest

Horizontal Gene Transfer

Assessment of genes associated with horizontal gene transfer (GO: 0009292) (Figure 4.8) revealed statistically significant (Wilcoxon, p-value < 0.05) differences in 16S rRNA normalized gene counts between processes upstream of ozonation and BAF downstream of ozonation. Specifically, both BAC filtration and GAC contacting, regardless of EBCT, were found to house elevated relative abundances of horizontal gene transfer related genes compared to influent and FlocSed. Not only were relative gene abundances higher during BAC filtration and GAC contacting, but gene profiles clustered by location (Bray Curtis; ANOSIM, p-value < 0.05, r-stat = 0.341) (SI Figure A.1). The relative abundance of HGT associated genes within ozonation were also correlated with nitrite adjusted O3:TOC ratios (Pearson, p-value < 0.10, r = 0.68).

BAC effluents, regardless of EBCT, were also found to be negatively correlated with the sample's proximity to backwashing (Spearman, p-value < 0.05, rho = -0.53).

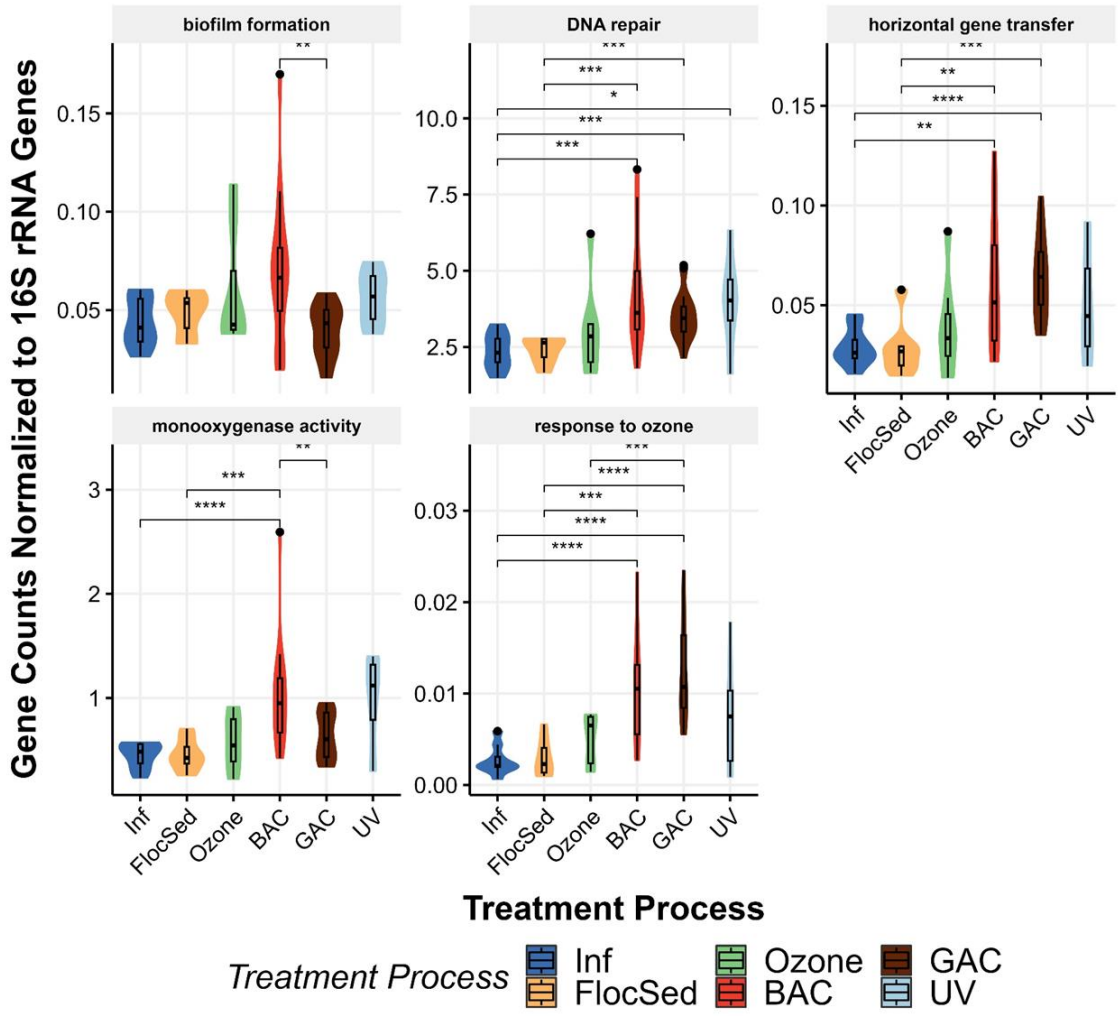


Figure 4.8: Functional gene abundances normalized by 16S rRNA gene abundance for selected functions of interest and grouped by treatment location. Statistically significant differences between sample locations were determined via Wilcox ranked sum test with an adjusted cutoff of 0.05. Selected functions of interest include horizontal gene transfer, biofilm formation, monooxygenase activity, DNA repair, and response to ozone.

Biofilm Formation

Genes associated with biofilm formation (GO:0042710) (Figure 4.8) were found to be significantly higher within the BAC filters compared to GAC contacting (Wilcoxon, p-value<0.05). Both disinfection processes, ozonation and UV, also experienced periods of elevated relative abundances, but were not found to be statistically significant. Biofilm gene profiles were also found to be clustered by treatment processes prior to BAC filtration and after BAC filtration (ANOSIM, p-value < 0.05, r-stat = 0.41), with ozonation seemingly providing a

shift in the makeup of genes associated with biofilm formation. When assessing variability within each treatment process (SI Figure A.2), BAC5 experienced a noticeable uptick in biofilm related genes following the addition of filter aid polymer (S16 and S17). This phenomenon was not observed in the higher EBCT BAC10. Both BACs also experienced lower relative abundances of biofilm formation genes when chlorine quenching was completed using sodium thiosulfate, except for S12, when temperature control began. UV disinfection experienced lowest biofilm related relative abundances during S14, S16, and S17 which also corresponded with the functional gene profiles that were most divergent from GAC effluents on multiple NMDS plots spanning full categorization of biological processes, to selected gene categories of interest (e.g. horizontal gene transfer, biofilm). Similarly, to UV disinfection, biofilm related genes coupled with ozone dosing seemed to have an impact of the overall functional profiles of ozone effluents. Specifically, samples with higher biofilm gene abundances (S05 and S10) and/or higher O3:TOC ratios (S03, S10, and S11) resulted in the greatest shifts in functional profiles while events with average biofilm gene abundances and lower ozone doses resulted in minimal shifts (S13, S14, and S16).

Response to Ozone and DNA Repair

Following ozone disinfection both BAC filters and GAC contactors experienced enriched relative abundances (Wilcoxon, p-value<0.05) of genes related to ozone response (GO:0010193)(Figure 4.8). Compared to the influent and FlocSed effluent, ozone response genes were found in slightly elevated relative abundances directly after ozone with statistically significant elevation after BAC and GAC treatment where DO concentrations stayed high. These enrichments peaked within BAC10 and GAC10, with relative gene abundances trending back to normal by UV disinfection. Further, the lowest relative abundances of response genes within ozone effluents were associated with the lowest adjusted O3:TOC ratios while GAC effluent abundances, regardless of EBCT, were strongly correlated with adjusted O3:TOC ratios (Pearson, p-value<0.05, r = 0.59). Similarly to ozone response, assessment of relative gene abundances associated with DNA repair (GO:0006281) saw slight enrichments directly after ozonation with statistically significant enrichment during BAC and GAC contacting (Wilcoxon, p-value<0.05). Statistically significant enrichment was also identified between processes prior to any exposed to disinfection and UV effluent (Wilcoxon, p-value<0.05) which was also slightly elevated from GAC contacting.

Monooxygenase Activity

Functional gene profiles related to monooxygenase activity (GO:0004497) were found to associated with individual treatment processes (ANOSIM, p-value<0.05, r-stat = 0.47) (SI Figure A.3). These functional profiles were notably different between front end processes prior to ozonation with well separated clusters between BAC filtration and GAC contacting. When assessing the relative abundance of monooxygenase related functional genes, a similar distinction was made with the BAC filters experiencing significantly enriched abundances when compared to influent, FlocSed, and GAC contacting (Wilcoxon, p-value<0.05), Figure 4.8. Both ozonation and GAC contacting were also slightly enriched when compared to influent and Flocsed, however, these observations were not statistically significant. Correlation analysis revealed that no statistically significant relationships were observed between genes related to monooxygenase activity in BAC5 and water quality data, while this relationship in BAC10 a negative correlation with influent TOC concentrations (Spearman, p-value<0.05, rho = -0.92).

Further, both GAC10 ((Spearman, p-value<0.05, rho = -0.82) and GAC20 (Spearman, p-value<0.05, rho = -0.89) were found to be negatively correlated with influent total inorganic nitrgoen. When considering relative gene abundances associated with all oxidoreductase activities (GO:0016491) enrichment was found to be statistically significant within both BAC filtration and GAC contacting when compared to influent and FlocSed (Wilcoxon, p-value<0.05).

Nitrogen Cycling

Overall, gene profiles related to nitrogen cycling were markedly distinct between process prior to ozonation and after ozonation (ANOSIM, p-value<0.05, r-stat = 0.45) (SI Figure A.4). Genes related to nitrogen cycling were grouped into 8 categories, specifically: ammonia monooxygenase activity (GO:0018597), nitrate assimilation (GO:0042128), nitrate reductase activity (GO:0008940), nitric oxide reductase activity (GO:0016966), nitrite reductase activity (GO:0098809), nitrogen fixation (GO:0009399), nitrous-oxide reductase activity (GO:0050304), and urease activity (GO:0009039). Of these, relative abundances were found to best highest within the urease activity, nitrogen fixation, nitrate assimilation, and nitrate reductase activity categories, (Figure 4.9). The categories associated with monooxygenase activity, nitric oxide reductase activity, nitrous-oxide reductase activity, and nitrite reductase activity were found in slightly lower relative abundances.

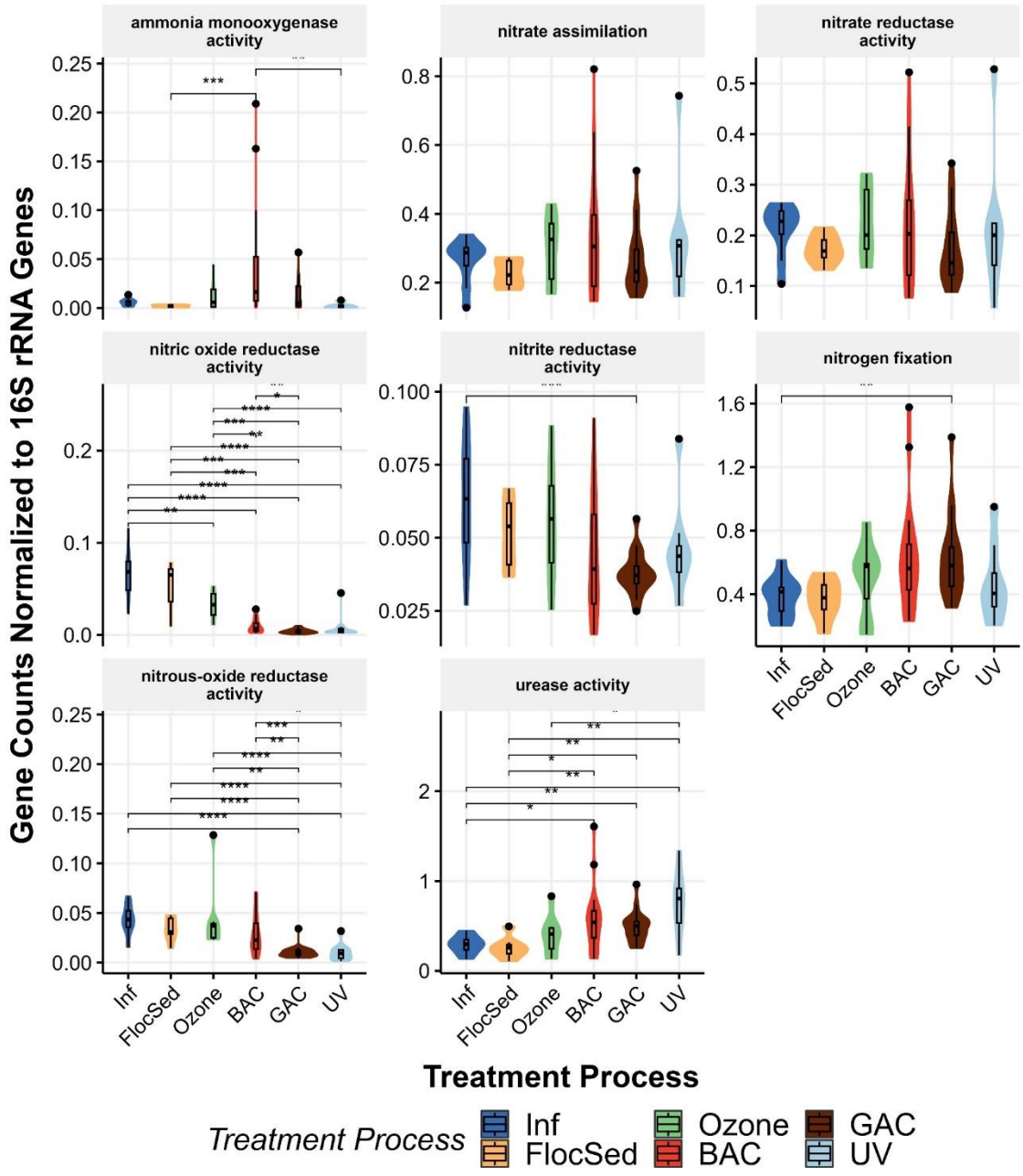


Figure 4.9: Functional gene abundances normalized by 16S rRNA gene abundance for selected functions of interest related to nitrogen cycling and grouped by treatment location. Statistically significant differences between sample locations were determined via Wilcox ranked sum test with an adjusted cutoff of 0.05. Selected functions related to nitrogen cycling included ammonia monooxygenase activity (GO:0018597), nitrate assimilation (GO:0042128), nitrate reductase activity (GO:0008940), nitric oxide reductase activity (GO:0016966), nitrite reductase activity (GO:0098809), nitrogen fixation (GO:0009399), nitrous-oxide reductase activity (GO:0050304), and urease activity (GO:0009039).

No statistically significant differences were found between treatment processes with respect to the nitrate assimilation or nitrate reductase activity categories. Genes related to ammonia monooxygenase activity were found to be close to zero at the influent and effluent of the pilot , with statistically significant enrichment during BAC filtration (Wilcoxon, p-value <0.05) and tapered relative abundances both before and after. Nitric oxide reductase activity was found to be statistically higher within the influent and FlocSed effluents (Wilcoxon, p-value <0.05), with moderate reductions during ozonation relative to influent (Wilcoxon, p-value <0.05). BAC filtration provided the greatest loss of related genes (Wilcoxon, p-value <0.05), with relative abundances being the lowest during GAC contacting and UV effluent. A similar relationship was seen with the genes associated with nitrous-oxide reductase activities where relative abundances persisted in higher relative abundances through ozonation with only moderate, non-significant reductions within BAC effluents. Statistically significant reductions in gene abundances, relative to influent abundances, were observed within the GAC contactors and UV effluents (Wilcoxon, p-value <0.05). Genes related to nitrite reductase activity experienced limited reductions in relative abundances through ozonation, with slightly greater reductions during BAC treatment and statistically significant reductions during GAC contacting (Wilcoxon, p-value <0.05), relative to influent abundances. Conversely, genes associated with nitrogen fixation were found to experience relatively low abundances early in the treatment train with significant enrichment occurring during GAC contacting (Wilcoxon, p-value <0.05), relative to influent, with slightly lower relative abundances during ozonation and BAC filtration. Finally, urease activity related relative abundances were found to be significantly higher within BAC filtration, GAC contacting, and UV disinfection relative to treatments prior to ozonation (Wilcoxon, p-value <0.05).

DISCUSSION

This study provides a comprehensive assessment of the functional gene dynamics characteristic throughout a carbon-based advanced treatment train intended for potable reuse. This work combined high level functional characterizations and more targeted assessment of functional genes of interest over a range of operational conditions and water quality parameters. Overall, the results revealed wide-ranging impact of treatment processes on the functional gene profile and the ability for shotgun metagenomic sequencing to shed light on these molecular dynamics. Specifically, the functional profiles identified within collected samples were highly consistent with expectations as a function of their specific treatment processes. Microbial communities clearly filled key environmental niches, in the case of biological filtration, or respond to shock conditions, in the case of microbial volatile disinfection. Interestingly, in most comparisons very little difference was seen between the influent and FlocSed effluent, indicating that physical separation alone was not enough to impart appreciable change on functional, metabolic, or other targeted functional profiles. Ozonation, in many comparisons, was found to provide a variable degree of change, ranging from small to large depending on the comparison of interest and specific operational parameters. Most notably, nitrite-adjusted O3:TOC ratio, which also impacted the functional profiles within BAC filtration and GAC contacting. Limited replication at higher ozone doses coupled with several lower tested doses resulted in fewer statistically significant relationships. The functional profiles in the treatment train influent and effluent were also notably different in several conducted analysis, highlighting the combined impact of the applied treatments on significantly altering the functional makeup of the effluent communities when compared to the influent.

When considering the entire functional and metabolic profiles, it was notable that BAC filtration and GAC contacting were functionally distinct processes. On the other hand, EBCT resulted in little distinction in functional gene profiles within each of these unit processes. A key finding of this research is that the GAC contactors were rich in functional genes, and thus as the GAC ages beyond its sorpative capacity, they essentially evolve as an additional biological treatment step. This research further demonstrates that the functional gene profiles of GAC contactors were distinct than preceding BAC filters, indicating that the corresponding dominant microbial functions occurring are likely distinct. It is hypothesized from this work (Chapter 5) and others[38–40], that this distinction is related to differences in the organic makeup of the process inlet, different nutrient compositions, and operational considerations. It is also worth pointing out that GAC contactors are backwashed much less frequently than BAC filters, leading to greater degrees of stratification with bed depth and the ability to develop more specialized metabolic pathways. Additionally, GAC contactors are protected from the highly oxidizing conditions occurring directly after ozonation and any supplemental chemical additions prior to BAC filtration (e.g. filter aid polymer, nutrients, monochloramine, chlorine quenching agents), which had notably more impact on BAC than GAC. Conversely, both BAC and GAC treatments were tied to pilot runtime more so than other treatment process, which is consistent with influence of temporal dimensions, such as succession, on the development of the microbial communities involved in biological treatment and their overall tendency to reflect and retain historic influences from changing operation conditions, relative to physical or chemical treatment stages.

When annotated functional genes were categorized according to high level biological processes, significant enrichment within functions related to biological activity, metabolic processes, and reproduction were observed after ozonation during BAC filtration and GAC contacting. This confirms the elevated biological removal occurring during BAF is supported by increased relative gene abundances. Further, it confirms the ability for ozonation to stimulate biological degradation in treated effluents subjected to multiple stages of biological treatment upstream and for metagenomic sequencing to capture these changes. Genes related to reproduction following disinfection and during BAC operation and GAC colonization highlight the ability for functional metagenomics to shed light on regrowth after disinfection[41–44] and microbial succession during biological treatment. Interestingly, detoxification abundances were highest during BAC filtration and UV effluent, consistent with response of microorganisms to microbial toxic conditions. The fact that such toxic-response genes remained elevated (and variable) during BAC filtration supports the hypothesis that microbially toxic conditions developed during ozonation carry over into the BAC filters[4]. The removal of genes related to viral process following ozonation with continued attenuation during the following processes supports the conclusion that midstream ozone disinfection does contribute to reduction of viral loads, without regrowth, while also supporting NGS technology's ability to identify trends in less abundant categories of interest.

Trends identified within the high-level metabolic categories were generally like those associated with the high-level classifications of biological processes, with a few notable highlights. Generally, higher relative abundances of metabolic related genes were found in BAC filters than GAC contactors, which were typically on par with ozone effluents. This is consistent with the intended function of the BAC filters, but was surprising in that, after prolonged operation (which most of the GAC studies collected in this study were subjected to), GAC contactors would have been expected to more closely reflect the BAC filters, as they evolve to

play more of an active role in biological removal[5,45]. Antibiotic metabolic processes (genes related to both antibiotic resistance and degradation) also being higher during BAC filtration and GAC contacting relative to the pilot influent, especially at greater EBCT, suggests that compounds sorbed on GAC media, including antibiotics or other CECs, remain bioavailable and could exert selection pressure on microbial communities. Potential antibiotic resistance selection pressures are expanded upon in Chapter 3. Lower EBCT also generally resulted in greater variability in identified relative abundances, speaking to potential process volatility with the BAC filters, especially within BAC5, generally displaying greater variance in functional gene composition than the GAC contactors. Overall, trends within extremely high-level categories were not entirely expected, as genes corresponding to any of these broad categories are fundamental to survival. The fact that differences exist between treatment process, even at these high levels, implies that functional gene profiling via metagenomics does in fact hold promise as a diagnostic tool for potable reuse applications[46].

Assessment of differentially abundant genes supported the conclusion that functional gene makeup between influent and FlocSed was very consistent, as more genes were identified relative to other treatment processes when they were combined than separated. The highly oxidizing nature of ozonation and wide range of doses applied during this study period lead to few consistently differentially abundant genes relative to other treatment processes, which is consistent with ozone to indiscriminately destroy DNA and other cell components. Interestingly, the differential abundance analysis and identification of corresponding functional genes resulted in divergent pictures of the BAC and GAC treatments' functional profiles. As stated earlier, the full characterization of functional genes into metabolic processes commonly found increased relative abundances present within the BAC filters while differential abundance analysis identified significantly more metabolically related functionals genes during GAC contacting. These distinct results, while seemingly at odds, shed light on interesting dynamics present within both treatment processes when operated consecutively. BAC filters, following ozonation, have access to increased levels of bioavailable carbon[47–49] (Chapter 2) and are routinely backwashed, which prevents stratification of degradation zones[50]. The combination of which limits selective pressure for more complex metabolic pathways, or niche organisms (Chapter 5), and results in the utilization of a wide range of more conventional degradative genes. On the other hand, GAC contacting provides the selective pressure for more complex metabolic pathways and diverse organisms resulting in the presence of significantly more differentially abundance functional genes related to metabolic processes. These dynamics were realized in a study that found BAC filters typically removed very few CECs, likely preoccupied with the degradation of bioavailable carbon and degraded CECs generated during ozonation, compared to GAC contactors, even after GAC exaustion[23].

Assessment of the fate of differential genes directly between treatments (not compared to all other treatments) indicated that, while UV disinfection resulted in a high number of differentially abundant genes when compared to GAC effluent, it failed to introduce unique genes when compared to other processes. This is consistent with the understanding that UV disinfection imparts significant alteration of gene abundances instead of selection of novel genes. Interestingly, there was striking overlap between dominant functional genes profiled post ozonation and post UV disinfection, highlighting commonalities in how these disinfection processes functionally impact microbial communities. BAC5's lack of unique genes following ozonation may be a result of its higher susceptibility to variance in operating conditions and

correspondingly diverse range of functional genes that are enriched with time, while GAC10 experienced the second highest counts of differentially abundant genes, speaking to its relative stability following BAC5. Further, BAC5 experienced a much greater percentage of its identified differentially abundance genes to be shared with other treatment process, indicating that it failed to provide many unique functional niches relative to other treatment process. Specifically, the greatest number of genes were shared between BAC and GAC treatments operated at different EBCT (BAC5|BAC10 and GAC10|GAC20) with fewer shared genes between BAF treatments. This supports the overarching conclusion that greater functional differences exist between BAC and GAC treatments than within either process operated at different EBCT. Interestingly, BAC10 did experience significantly more shared, differential genes with the GACs than BAC5, indicating that longer EBCT can somewhat bridge the functional gap experienced between the two systems. Overall, differential abundance analysis was found to be effective at identifying uniquely abundant functional genes present within different stages of treatment, especially those related to biological treatment.

To obtain insight into higher resolution dynamics presented within individual treatments, more targeted analysis was conducted on specific functions of interest, including: horizontal gene transfer, biofilm formation, monooxygenase activity, response to disinfection, and nitrogen cycling. Horizontal gene transfer is of particular importance due to its ability for organisms to share beneficial genes, improving biological processes and microbial community dynamics[51–55], or genes associated with concerns for public health[56–58]. Here we found that genes associated with horizontal gene transfer were most enriched within the complex microbial communities associated with BAC filtration and GAC contacting, while genes within ozonation were found to have highest relative abundances when oxidative stress was highest during periods of elevated ozone dosing. Interestingly, relative abundances of horizontal transfer related genes were negatively correlated with their temporal proximity to backwashing, supporting the hypothesis that horizontal gene transfer is heighted directly after media is redistributed during backwashing and microbial communities are exposed to new environmental niches within the biofilter.

Biofilms were also of particular interest both before, during, and after biological filtration, due to their ability to stimulate complex microbial interactions[52,59], increase degradative capacity[52,60–62], promote fouling[63], and protect against disinfection[64]. Within this study, genes related to biofilm formation were found to be highest during BAC filtration where increased bioavailable carbon coupled with disruptive backwashing could create an environmental of constant microbial growth and biofilm turnover. Relative gene abundances were also highest within BAC5 during the addition of filter aid polymer, which could act as a scaffold, facilitating increased biofilm formation. The same phenomenon was not experienced within BAC10, where the higher EBCT resulted in a lower mass loading of filter aid polymer onto the filter and a less frequent need for backwashing.. Interestingly, biofilm formation genes on both BAC5 and BAC10 were lowest during periods of sodium thiosulfate quenching (compared to no quenching or quenching with sodium bisulfite), with a sharp increase midway through testing with the beginning of temperature control. When comparing the impact of biofilm formation on disinfection efficacy, it was found that samples with the lowest relative abundances of biofilm related genes experienced the largest shifts from reference GAC functional profiles. This makes sense as biofilms are known to be more resistant to UV disinfection[65–67] and the depletion of biofilms should allow for a more uniform removal of DNA and greater shift in the functional gene profile. During ozonation, the relative abundances of

biofilm-related functional genes and ozone dosing were related to the shifts in functional profiles. Specifically, sampling events with high relative abundances of biofilm related genes without relatively high ozone dosing resulted in the most divergent functional profiles, as determined via beta-diversity. This makes sense as biofilms should provide protection from ozonation, and influents with higher concentrations of biofilms should require elevated levels of ozone dosing to fully disinfection biofilm communities[68–70].

Functional genes associated with ozonation were found to be slightly higher directly after ozonation, relative to influent samples, and persisted during BAC filtration and GAC contacting before decreasing during UV disinfection. This highlights the important, long-lasting impact of ozonation on downstream microbial communities, where DO concentrations are still extremely high and can cause microbial stress[71]. Genes related to DNA repair experienced similar trends following ozonation and into BAC and GAC treatment, but were further enrichment following UV disinfection. Overall, the exposure and response nature of damage repair and stress response genes following disinfection makes sense and gives credibility to the application of metagenomic functional gene profiling in accessing informative changes within microbial functions. Similar observations were seen with monooxygenase activities, known for their ability to provide CEC removal[72–75] and less discriminatory degradation[76–78], increasing during BAC filtration.

Metagenomic functional gene profiling was also found to be effective at shedding light on the complex interplay between various forms of inorganic nitrogen throughout the treatment train. Nitrogen is known for its paramount importance within biological processes and is often a growth-limiting nutrient for important degradative processes.[9,79] Nitrogen is also a contaminant of its own right[80,81] and various treatment processes are applied to drive removal of its various redox states[82,83]. Genes encoding ammonia monooxygenase, a key enzyme in the transformation of ammonia during nitrification[84,85], were found to respond to the addition of preformed monochloramine prior to ozonation. Ammonia concentrations were generally found to be highest prior to BAC filtration, which was consistently effective in nitrifying ammonia to nitrate. Urease related genes were found to increase following ozonation, which was consistent with other studies noting elevated urease activity following ozone.[86–88] Urease can release ammonia through the hydrolysis of urea[89]. Denitrification was likely hindered in the BAC units due to elevated dissolved oxygen post ozone. Correspondingly, genes related to nitrate reductase activity, responsible for converting nitrate to nitrite[90] and present within many bacteria, were never found to be enriched in any treatment process. Likewise, nitrate was not observed to be substantially removed during treatment.

Interestingly, nitric oxide reductase genes were enriched to the greatest degree during influent processes and thereafter were attenuated throughout the rest of the treatment train. This is logical, given that as nitric oxide is an intermediate of denitrification pathways[91] and a dentification filter was operated directly upstream of the pilot influent. Increased nitric oxide reduction into nitrous oxide also resulted in enrichment of nitrous-oxide reductase genes, which are responsible for the further reduction of nitrous oxide to dinitrogen gas[92]. Nitrous-oxide reductase genes followed a similar enrichment profile to genes associated with nitric oxide reductase - albeit with a slight spatial delayed to its tapered reduction. Meanwhile, nitrite reductase genes were found to be highest where nitrite was found in appreciable concentrations, prior to GAC contacting. Nitrite reductase is more commonly known for its ability to reduce nitrite to ammonia and less commonly known to produce nitric oxide[93,94] making it an interesting

component to the nitrogen dynamics occurring earlier on in the treatment train when ammonia, nitrite, and nitric oxide coexisted. It is possible that while nitrite reductase activity was high, a portion of nitrite could have been denitrified via the nitric oxide – nitrous oxide – dinitrogen gas pathway ,contributing to the small overall reduction of total inorganic nitrogen found in BAC and GAC effluents.

Genes associated with nitrogen fixation, i.e., transformation of dinitrogen gas in the atmosphere into inorganic carbon[95–97], experienced a surprising uptick after exposure to ozonation within the BAC and GAC processes. This is somewhat unexpected, given that nitrogen fixation is an energy intensive process and thus traditionally was thought to occur only in symbiotic relationships, where a host organism provides organic carbon needed to sustain growth in exchange for the bioavailable nitrogen. However, free-living nitrogen-fixing bacteria have now been described and appear to be most active in environments where organic carbon is available, but nitrogen is limited.[98–100] Free-living nitrogen fixers, such as certain cyanobacteria, have been reported in water treatment systems, such as drinking water filters[101]. Interestingly, in the present study, the treatment processes were not open to the atmosphere, thus nitrogen gas likely penetrated the system via other means, such as via the dinitrogen gas supply used for the ozonation systems. Dinitrogen gas, coupled with bioavailable nitrogen limitation and freely available organic carbon within the BACs and GACs could have feasibly resulted in selection of organisms capable of nitrogen fixation. In particular, Cyanobacteria (SI Table A.1) were noted to increase in relative abundance following ozonation, relative to the influent, with high abundances persisting through the BAC filters and even higher abundances noted following GAC contacting, where nitrogen was the most limited.

Overall, the application of metagenomic functional gene profiling to assess genes within water treatment, wastewater treatment, and water reuse provided successful at shedding light on various mechanistic aspects of this multi-barrier carbon-based potable water reuse train. Differentially abundant gene analysis and tracking of specific categories of functional genes associated with functions of interest served to identify trends among abundant and less abundant biological processes, disentangle relationships within and between treatments, and characterize microbial community response and adaptation to changing operational conditions. Specifically, functional gene profiling proved able to illuminate interesting dynamics associated with metabolic process during biological treatment, interactions between disinfection and microbial response, and nitrogen cycling, including unexpected or less common metabolic pathways.

ACKNOWLEDGEMENTS

The authors would like to thank the operators and staff at Hampton Road Sanitation District's York River Treatment Plant and the SWIFT Research Center. Funding for this effort was provided in part by Water Research Foundation Unsolicited Award U1R16 (PI Pruden) and US Bureau of Reclamation grant R21AC10162 (PI Pruden), leveraged with additional financial and in-kind support from the Hampton Roads Sanitation District. Special thanks to Pruden Lab members for assistance in the laboratory.

REFERENCES

1. Liao, X. *et al.* Changes of biomass and bacterial communities in biological activated carbon filters for drinking water treatment. *Process Biochem.* **48**, 312–316 (2013).

2. Stoquart, C., Barbeau, B., Servais, P. & Vázquez-Rodríguez, G. A. Quantifying bacterial biomass fixed onto biological activated carbon (PAC and GAC) used in drinking water treatment. *J. Water Supply Res. Technol.* **63**, 1–11 (2014).

3. Liao, X. *et al.* Pyrosequencing analysis of bacterial communities in drinking water biofilters receiving influents of different types. *Process Biochem.* **48**, 703–707 (2013).

4. Lohwacharin, J., Phetrak, A., Takizawa, S., Kanisawa, Y. & Okabe, S. Bacterial growth during the start-up period of pilot-scale biological activated carbon filters: Effects of residual ozone and chlorine and backwash intervals. *Process Biochem.* **50**, 1640–1647 (2015).

5. Gibert, O. *et al.* Characterising biofilm development on granular activated carbon used for drinking water production. *Water Res.* **47**, 1101–1110 (2013).

6. Magic-Knezev, A. & van der Kooij, D. Optimisation and significance of ATP analysis for measuring active biomass in granular activated carbon filters used in water treatment. *Water Res.* **38**, 3971–3979 (2004).

7. Liu, X., Huck, P. M. & Slawson, R. M. Factors Affecting Drinking Water Biofiltration. *J. Am. Water Works Assoc.* **93**, 90–101 (2001).

8. Basu, O. D., Dhawan, S. & Black, K. Applications of biofiltration in drinking water treatment – a review. *J. Chem. Technol. Biotechnol.* **91**, 585–595 (2016).

9. Dhawan, S., Basu, O. D. & Banihashemi, B. Influence of nutrient supplementation on DOC removal in drinking water biofilters. *Water Supply* **17**, 422–432 (2017).

10. Ye, L., Zhang, T., Wang, T. & Fang, Z. Microbial structures, functions, and metabolic pathways in wastewater treatment bioreactors revealed using high-throughput sequencing. *Environ. Sci. Technol.* **46**, 13244–13252 (2012).

11. Cai, M. *et al.* Metagenomic Reconstruction of Key Anaerobic Digestion Pathways in Municipal Sludge and Industrial Wastewater Biogas-Producing Systems. *Front. Microbiol.* **7**, 778 (2016).

12. Sidhu, C., Vikram, S. & Pinnaka, A. K. Unraveling the Microbial Interactions and Metabolic Potentials in Pre- and Post-treated Sludge from a Wastewater Treatment Plant Using Metagenomic Studies. *Front. Microbiol.* **8**, 1382 (2017).

13. Folch-Mallol, J. L., Zárate, A., Sánchez-Reyes, A. & López-Lara, I. M. Expression, purification, and characterization of a metagenomic thioesterase from activated sludge involved in the degradation of acylCoA-derivatives. *Protein Expr. Purif.* **159**, 49–52 (2019).

14. Cai, P. *et al.* Insights into Biodegradation Related Metabolism in an Abnormally Low Dissolved Inorganic Carbon (DIC) Petroleum-Contaminated Aquifer by Metagenomics Analysis. *Microorganisms* **7**, 412 (2019).

15. Biderre-Petit, C., Taib, N., Gardon, H., Hochart, C. & Debroas, D. New insights into the

pelagic microorganisms involved in the methane cycle in the meromictic Lake Pavin through metagenomics. *FEMS Microbiol. Ecol.* **95**, 183 (2019).

16. Mohan, A. M., Bibby, K. J., Lipus, D., Hammack, R. W. & Gregory, K. B. The Functional Potential of Microbial Communities in Hydraulic Fracturing Source Water and Produced Water from Natural Gas Extraction Characterized by Metagenomic Sequencing. *PLoS One* **9**, e107682 (2014).

17. Chao, Y. *et al.* Metagenomic analysis reveals significant changes of microbial compositions and protective functions during drinking water treatment. *Sci. Reports 2013 31* **3**, 1–9 (2013).

18. Li, T., Guo, F., Lin, Y., Li, Y. & Wu, G. Metagenomic analysis of quorum sensing systems in activated sludge and membrane biofilm of a full-scale membrane bioreactor. *J. Water Process Eng.* **32**, 100952 (2019).

19. Douterelo, I., Calero-Preciado, C., Soria-Carrasco, V. & Boxall, J. B. Whole metagenome sequencing of chlorinated drinking water distribution systems. *Environ. Sci. Water Res. Technol.* **4**, 2080–2091 (2018).

20. Pruden, A., Bott, C., Blair, M. F., Miller, J. H. & Vaidya, R. Characterization of Organic Carbon and Microbial Communities for the Optimization of Biologically-Active Carbon (BAC) Filtration for Potable Reuse. *Water Res. Found. Proj. 4872/U1R16 Final Report.* 305-undefined (2020).

21. Sun, Y. *et al.* A pilot-scale investigation of disinfection by-product precursors and trace organic removal mechanisms in ozone-biologically activated carbon treatment for potable reuse. *Chemosphere* **210**, 539–549 (2018).

22. Vaidya, R. *et al.* Pilot Plant Performance Comparing Carbon-Based and Membrane-Based Potable Reuse Schemes. *Environ. Eng. Sci.* **36**, 1369–1378 (2019).

23. Vaidya, R., Wilson, C. A., Salazar-Benites, G., Pruden, A. & Bott, C. Implementing Ozone-BAC-GAC in potable reuse for removal of emerging contaminants. *AWWA Water Sci.* **2**, e1203 (2020).

24. Vaidya, R., Wilson, C. A., Salazar-Benites, G., Pruden, A. & Bott, C. Factors affecting removal of NDMA in an ozone-biofiltration process for water reuse. *Chemosphere* **264**, 128333 (2021).

25. Arango-Argoty, G. *et al.* MetaStorm: A Public Resource for Customizable Metagenomics Annotation. *PLoS One* **11**, (2016).

26. Bateman, A. *et al.* UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).

27. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).

28. Li, B. *et al.* Metagenomic and network analysis reveal wide distribution and co-

occurrence of environmental antibiotic resistance genes. *ISME J.* **9**, (2015).

29. R Core Team, 2018. R: A language and environment for statistical computing. *R Found. Stat. Comput. Vienna, Austria* (2018).

30. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25 (2000).

31. Carbon, S. *et al.* The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* **49**, D325–D334 (2021).

32. Carlson, M. A set of annotation maps describing the entire Gene Ontology. *R Packag. version 3.8.2* (2019).

33. Oksanen, J. *et al.* Package 'vegan' Title Community Ecology Package Version 2.5-7. (2020).

34. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, (2014).

35. Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R. & Pfister, H. UpSet: Visualization of intersecting sets. *IEEE Trans. Vis. Comput. Graph.* **20**, 1983–1992 (2014).

36. Harrell, F. E. Harrell Miscellaneous (Package ' Hmisc '). 421 (2018).

37. Wickham, H. ggpolt2 Elegant Graphics for Data Analysis. *Use R! Ser.* 211 (2016).

38. De Vera, G. A., Gerrity, D., Stoker, M., Frehner, W. & Wert, E. C. Impact of upstream chlorination on filter performance and microbial community structure of GAC and anthracite biofilters. *Environ. Sci. Water Res. Technol.* **4**, 1133–1144 (2018).

39. Guarin, T. C. & Pagilla, K. R. Microbial community in biofilters for water reuse applications: A critical review. *Sci. Total Environ.* **773**, 145655 (2021).

40. Nemani, V. A., McKie, M. J., Taylor-Edmonds, L. & Andrews, R. C. Impact of biofilter operation on microbial community structure and performance. *J. Water Process Eng.* **24**, 35–41 (2018).

41. Wang, M., Ateia, M., Awfa, D. & Yoshimura, C. Regrowth of bacteria after light-based disinfection — What we know and where we go from here. *Chemosphere* **268**, 128850 (2021).

42. Miettinen, I. T., Vartiainen, T., Nissinen, T., Tuhkanen, T. & Martikainen, P. J. MICROBIAL GROWTH IN DRINKING WATERS TREATED WITH OZONE, OZONE/HYDROGEN PEROXIDE OR CHLORINE. *http://dx.doi.org/10.1080/01919519808547266* **20**, 303–315 (2008).

43. Gorito, A. M. *et al.* Ozone-based water treatment (O3, O3/UV, O3/H2O2) for removal of organic micropollutants, bacteria inactivation and regrowth prevention. *J. Environ. Chem. Eng.* **9**, 105315 (2021).

44. de Vera, G. A. & Wert, E. C. Using discrete and online ATP measurements to evaluate regrowth potential following ozonation and (non)biological drinking water treatment. *Water Res.* **154**, 377–386 (2019).

45. Bancroft, K., Maloney, S. W., McElhaney, J., Suffet, I. H. & Pipes, W. O. Assessment of bacterial growth and total organic carbon removal on granular activated carbon contactors. *Appl. Environ. Microbiol.* **46**, 683–688 (1983).

46. Garner, E. *et al.* Next generation sequencing approaches to evaluate water and wastewater quality. *Water Res.* **194**, 116907 (2021).

47. Reungoat, J. *et al.* Ozonation and biological activated carbon filtration of wastewater treatment plant effluents. *Water Res.* **46**, 863–872 (2012).

48. Reungoat, J. *et al.* Removal of micropollutants and reduction of biological activity in a full scale reclamation plant using ozonation and activated carbon filtration. *Water Res.* **44**, 625–637 (2010).

49. Lehtola, M. J., Miettinen, I. T., Vartiainen, T., Myllykangas, T. & Martikainen, P. J. Microbially available organic carbon, phosphorus, and microbial growth in ozonated drinking water. *Water Res.* **35**, 1635–1640 (2001).

50. Liao, X. *et al.* Operational performance, biomass and microbial community structure: Impacts of backwashing on drinking water biofilter. *Environ. Sci. Pollut. Res.* **22**, 546–554 (2014).

51. French, Katherine E., Zhongrui Zhou, and N. T. Horizontal 'gene drives' harness indigenous bacteria for bioremediation. *Sci. Rep.* (2020).

52. Singh, R., Paul, D. & Jain, R. K. Biofilms: implications in bioremediation. *Trends Microbiol.* **14**, 389–397 (2006).

53. Bhandari, G. & Karn, S. K. Evaluation of horizontal gene transfer of catabolic genes and its application in bioremediation. *Smart Bioremediation Technol. Microb. Enzym.* 359–372 (2019) doi:10.1016/B978-0-12-818307-6.00019-6.

54. Shahi, A., Aydin, S., Ince, B. & Ince, O. Evaluation of microbial population and functional genes during the bioremediation of petroleum-contaminated soil as an effective monitoring approach. *Ecotoxicol. Environ. Saf.* **125**, 153–160 (2016).

55. Mishra, A. P., Sahoo, J. P., Padhi, P. P., Pattnaik, S. & Jena, L. Applications of Horizontal Gene Transfer in Soil Bioremediation. *Bioremediation Phytoremediation Technol. Sustain. Soil Manag.* 171–203 (2022) doi:10.1201/9781003281177-10.

56. Lindsay, J. A. Staphylococcus aureus genomics and the impact of horizontal gene transfer. *Int. J. Med. Microbiol.* **304**, 103–109 (2014).

57. Boto, L., Pineda, M. & Pineda, R. Potential impacts of horizontal gene transfer on human health and physiology and how anthropogenic activity can affect it. *FEBS J.* **286**, 3959–3967 (2019).

58. Miller, J. H., Novak, J. T., Knocke, W. R. & Pruden, A. Survival of antibiotic resistant bacteria and horizontal gene transfer control antibiotic resistance gene content in anaerobic digesters. *Front. Microbiol.* **7**, 263 (2016).

59. Moons, P., Michiels, C. W. & Aertsen, A. Bacterial interactions in biofilms. *http://dx.doi.org/10.1080/10408410902809431* **35**, 157–168 (2009).

60. Sonawane, J. M., Rai, A. K., Sharma, M., Tripathi, M. & Prasad, R. Microbial biofilms: Recent advances and progress in environmental bioremediation. *Sci. Total Environ.* **824**, 153843 (2022).

61. Mangwani, N., Kumari, S. & Das, S. Bacterial biofilms and quorum sensing: fidelity in bioremediation technology. *http://dx.doi.org/10.1080/02648725.2016.1196554* **32**, 43–73 (2016).

62. Sharma, A., Jamali, H., Vaishnav, A., Giri, B. S. & Srivastava, A. K. Microbial biofilm: An advanced eco-friendly approach for bioremediation. *New Futur. Dev. Microb. Biotechnol. Bioeng. Microb. Biofilms Curr. Res. Futur. Trends Microb. Biofilms* 205–219 (2020) doi:10.1016/B978-0-444-64279-0.00015-3.

63. Melo, L. F. & Bott, T. R. Biofouling in water systems. *Exp. Therm. Fluid Sci.* **14**, 375–381 (1997).

64. Percival, S. L. & Walker, J. T. Potable water and biofilms: A review of the public health implications. *http://dx.doi.org/10.1080/08927019909378402* **14**, 99–115 (2009).

65. de Carvalho, C. C. C. R. Biofilms: Microbial strategies for surviving UV exposure. *Adv. Exp. Med. Biol.* **996**, 233–239 (2017).

66. Yin, W., Wang, Y., Liu, L. & He, J. Biofilms: The Microbial "Protective Clothing" in Extreme Environments. *Int. J. Mol. Sci. 2019, Vol. 20, Page 3423* **20**, 3423 (2019).

67. Elasri, M. O. & Miller, R. V. Study of the response of a biofilm bacterial community to UV radiation. *Appl. Environ. Microbiol.* **65**, 2025–2031 (1999).

68. Kotlarz, N. *et al.* Biofilms in Full-Scale Drinking Water Ozone Contactors Contribute Viable Bacteria to Ozonated Water. *Environ. Sci. Technol.* **52**, 2618–2628 (2018).

69. Lund, V. & Ormerod, K. The influence of disinfection processes on biofilm formation in water distribution systems. *Water Res.* **29**, 1013–1021 (1995).

70. Momba, M. N. B., Cloete, T. E., Venter, S. N. & Kfir, R. Evaluation of the impact of disinfection processes on the formation of biofilms in potable surface water distribution systems. *Water Sci. Technol.* **38**, 283–289 (1998).

71. Tikariha, H., Khardenavis, A. A. & Purohit, H. J. Dissolved oxygen-mediated enrichment of quorum-sensing phenomenon in the bacterial community to combat oxidative stress. *Arch. Microbiol.* **200**, 1371–1379 (2018).

72. Mahendra, S. & Alvarez-Cohen, L. Kinetics of 1,4-Dioxane Biodegradation by

Monooxygenase-Expressing Bacteria. *Environ. Sci. Technol.* **40**, 5435–5442 (2006).

73. Mahendra, S., Petzold, C. J., Baidoo, E. E., Keasling, J. D. & Alvarez-Cohen, L. Identification of the Intermediates of in Vivo Oxidation of 1,4-Dioxane by Monooxygenase-Containing Bacteria. *Environ. Sci. Technol.* **41**, 7330–7336 (2007).

74. Li, M. *et al.* The Abundance of Tetrahydrofuran/Dioxane Monooxygenase Genes (thmA/dxmA) and 1,4-Dioxane Degradation Activity Are Significantly Correlated at Various Impacted Aquifers. *Environ. Sci. Technol. Lett.* **1**, 122–127 (2013).

75. Deng, D., Li, F. & Li, M. A Novel Propane Monooxygenase Initiating Degradation of 1,4-Dioxane by Mycobacterium dioxanotrophicus PH-06. *Environ. Sci. Technol. Lett.* **5**, 86–91 (2018).

76. Bruce E. Rittmann, Eric Seagren, B. A. . W. In Situ Bioremediation. https://books.google.com/books?hl=en&lr=&id=7R7VX9dLV8oC&oi=fnd&pg=PA1&dq=in+situ+bioremediation+bruce+rittman&ots=fIOgdu4L-E&sig=-LrxwQ6cUTzPAe78eks1wVx_t-w#v=onepage&q=in situ bioremediation bruce rittman&f=false.

77. Harayama, S. Polycyclic aromatic hydrocarbon bioremediation design. *Curr. Opin. Biotechnol.* **8**, 268–273 (1997).

78. Okino-Delgado, C. H., Zanutto-Elgui, M. R., do Prado, D. Z., Pereira, M. S. & Fleuri, L. F. Enzymatic Bioremediation: Current Status, Challenges of Obtaining Process, and Applications. 79–101 (2019) doi:10.1007/978-981-13-7462-3_4.

79. Liao, X. *et al.* Biomass development in GAC columns receiving influents with different levels of nutrients. *Water Supply* **16**, 1024–1032 (2016).

80. O'Connor, G. A., Elliott, H. A. & Bastian, R. K. Degraded Water Reuse: An Overview. *J. Environ. Qual.* **37**, S-157 (2008).

81. Wiesmann, U. Biological nitrogen removal from wastewater. *Adv. Biochem. Eng. Biotechnol.* **51**, 113–154 (1994).

82. Lim, S., Shi, J. L., von Gunten, U. & McCurry, D. L. Ozonation of organic compounds in water and wastewater: A critical review. *Water Res.* **213**, 118053 (2022).

83. Sörensen, H. & Jörgensen, S. E. *The Removal of Nitrogen Compounds from Wastewater*. (1993).

84. Hooper, A. B., Vannelli, T., Bergmann, D. J. & Arciero, D. M. Enzymology of the oxidation of ammonia to nitrite by bacteria. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **71**, 59–67 (1997).

85. Rotthauwe, J. H., Witzel, K. P. & Liesack, W. The ammonia monooxygenase structural gene amoA as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing populations. *Appl. Environ. Microbiol.* **63**, 4704–4712 (1997).

86. Aparicio, M. A., Eiroa, M., Kennes, C. & Veiga, M. C. Combined post-ozonation and biological treatment of recalcitrant wastewater from a resin-producing factory. *J. Hazard. Mater.* **143**, 285–290 (2007).

87. Muhammad Yusuf, A. R., Mulana, F. & Said, S. D. Effects of ultraviolet-enhanced ozonation on the degradation of ammonia and urea in fertilizer plant wastewater. *IOP Conf. Ser. Mater. Sci. Eng.* **536**, 012079 (2019).

88. Benitez, F. J., Real, F. J., Acero, J. L. & Garcia, C. Kinetics of the transformation of phenyl-urea herbicides during ozonation of natural waters: Rate constants and model predictions. *Water Res.* **41**, 4073–4084 (2007).

89. von Ahnen, M., Pedersen, L. F., Pedersen, P. B. & Dalsgaard, J. Degradation of urea, ammonia and nitrite in moving bed biofilters operated at different feed loadings. *Aquac. Eng.* **69**, 50–59 (2015).

90. Park, J. Y. & Yoo, Y. J. Biological nitrate removal in industrial wastewater treatment: which electron donor we can choose. *Appl. Microbiol. Biotechnol. 2009 823* **82**, 415–429 (2009).

91. Fuerhacker, M. *et al.* Relationship between release of nitric oxide and CO2 and their dependence on oxidation reduction potential in wastewater treatment. *Chemosphere* **44**, 1213–1221 (2001).

92. Pauleta, S. R., Dell'Acqua, S. & Moura, I. Nitrous oxide reductase. *Coord. Chem. Rev.* **257**, 332–349 (2013).

93. Cabello, P., Roldán, M. D., Castillo, F. & Moreno-Vivián, C. Nitrogen Cycle. *Encycl. Microbiol.* 299–321 (2009) doi:10.1016/B978-012373944-5.00055-9.

94. Losada, M. & Paneque, A. Nitrite Reductase. *Methods Enzymol.* **23**, 487–491 (1971).

95. Dixon, R. & Kahn, D. Genetic regulation of biological nitrogen fixation. *Nat. Rev. Microbiol. 2004 28* **2**, 621–631 (2004).

96. John Postgate. *Nitrogen Fixation*. (1998).

97. Zehr, J. P. & Turner, P. J. Nitrogen fixation: Nitrogenase genes and gene expression. *Methods Microbiol.* **30**, 271–286 (2001).

98. Dynarski, K. A. & Houlton, B. Z. Nutrient limitation of terrestrial free-living nitrogen fixation. *New Phytol.* **217**, 1050–1061 (2018).

99. Ospina-Betancourth, C. *et al.* Enrichment of Nitrogen-Fixing Bacteria in a Nitrogen-Deficient Wastewater Treatment System. *Environ. Sci. Technol.* **54**, 3539–3548 (2020).

100. Pratt, S., Tan, M., Gapes, D. & Shilton, A. Development and examination of a granular nitrogen-fixing wastewater treatment system. *Process Biochem.* **42**, 863–872 (2007).

101. Gerrity, D. *et al.* Microbial community characterization of ozone-biofiltration systems in drinking water and potable reuse applications. *Water Res.* **135**, 207–219 (2018).

CHAPTER 5: RELATING MICROBIAL COMMUNITY COMPOSITION AND DYNAMICS TO TREATMENT PERFORMANCE IN AN ADVANCED OXIDATION PROCESS-BIOLOGICALLY ACTIVE CARBON FILTRATION-DRIVEN POTABLE REUSE TREATMENT TRAIN

Matthew F. Blair, Ramola Vaidya, Germano Salazar-Benites, Charles Bott, and Amy Pruden

ABSTRACT

Water stress and associated issues have led toward a push for more sustainable applications of advanced treatment trains applied for water reuse. Treatment trains that couple advanced oxidation processes (AOPs), such as ozone, with biologically active filtration (BAF) are of interest as a lower cost, membrane-free technology. However, little is known about the microbial communities that are the fundamental drivers of AOP-BAF treatment. The overarching objective of this study was to demonstrate the potential for microbial community profiling as a diagnostic tool for assessing the functionality and resilience of sequential processes employed in a water reuse treatment train. We utilized 16S rRNA gene amplicon sequencing to profile the bacterial microbiota characteristic of each stage of treatment with time in a carbon-based potable reuse train employing coagulation, flocculation, sedimentation, ozonation, BAC filtration, GAC contacting, and UV disinfection. It was found that there was a distinct baseline microbiota associated with each stage of treatment, each of which undergoing succession with time. Oxidative treatments, such as O_3, resulted in the sharpest shifts, and also variance, in microbial community composition. Other operational changes, such as adjustment in O_3:TOC, temperature, filter aid polymer, chlorine quenching agent, and EBCT also resulted in measurable changes in the baseline microbial community composition of each process, but to a lesser degree. Of these, supplementation of nitrogen and phosphorus resulted in the strongest bifurcation and was particularly noted in the analysis microbial communities inhabiting the BAC and GAC units. The findings of this study improve understanding of bacterial dynamics occurring in advanced water treatment trains and help to factor this understanding into improved system design and operation.

INTRODUCTION

Water scarcity coupled with the ecological impact of conventional wastewater treatment discharges underscore the need for a paradigm shift in the water industry toward developing, understanding, and implementing treatment trains designed for water reuse applications. Water reuse applications that augment conventional potable water sources with advanced tertiary treatment of secondary wastewater effluents are a promising means to enhance water supplies, while also addressing issues such as ground-level subsidence, sea water intrusion, and eutrophication of receiving water bodies. Cost-effective tertiary treatment trains that can reliably eliminate associated health risks and environmental hazards of wastewater effluents; specifically pathogens, suspended solids, nutrients, organic pollutants, and contaminants of emerging concern (CECs), is essential for expanding potable reuse application [1].

In many parts of the world, the default means of achieving high quality water for reuse purposes is through membrane filtration; such as reverse osmosis (RO) and ultrafiltration (UF),

which has significant drawbacks in terms of elevated costs, operational complexities, and the production of difficult to dispose of waste streams. An alternative to membrane filtration is the use of biologically-active filtration (BAF) which can take a number of forms, including biologically-active carbon (BAC) filtration. In BAC treatment, granular activated carbon (GAC) provides synergistic advantages of high surface area for adsorption and microbial colonization, which can further drive degradation of organics and uptake or transformation of nutrients. Further, ozone pre-treatment, a widely used advanced oxidative processes (AOP), can be applied to achieve pathogen removal as well as to enhance bioavailability of refractory organic compounds for biodegradation in a subsequent BAF/BAC unit. Implementation of BAC with preliminary ozone conditioning has been reported to achieve similar quality effluents as RO and UF systems with respect to trace organic carbon composition, but at a much lower cost [2]. Thus, such "carbon-based" treatment trains are a promising alternative to membrane treatment, but a fundamental understanding of the ultimate range of achievable water quality is needed to advance its full-scale application. In particular, there is a lack of understanding of the role of the microbial community in metabolizing organic matter and the factors governing the microbiological and organic content of final effluents.

Addressing the above knowledge gap is challenging from a methodological standpoint. Conventional microbiological monitoring of drinking water and wastewater treatment processes relies on culture-based methods, with regulatory requirements generally focusing on fecal indicator bacteria. However, culture-based methods are severely hampered when the aim is to characterize the broader microbial community composition, given the diversity and lack of culturability of the vast majority of microbes inhabiting these systems. Next generation DNA sequencing (NGS) technologies hold promise for both research and practical application in the water sector [3]. In particular, NGS techniques can help improve understanding of complex microbial dynamics and interactions that underlie full-scale biological treatment processes through their ability to rapidly recover millions of DNA sequences from complex environmental samples [4].

Given that microbial communities are fundamental drivers of AOP-BAF treatment, better understanding their composition and function can potentially yield new insight into treatment performance and interactions between coupled treatment processes. Of particular interest is the nature of microbes that initially colonize BAF units and key attributes of their successional patterns, with time and in response to changing environmental and operational conditions. In the case of AOP-BAC-GAC, there is a lack of understanding how the highly antimicrobial upstream oxidative environment created by AOPs, such as O_3, influences the microbes that flourish in downstream BAC filters and their overall performance. Further, while GAC treatment is intended to be driven primarily by abiotic processes, such as sorption, the reality is that GAC filters are a rich environmental niche for a wide diversity of microbes [5–7]. The extent to which microbes colonizing GAC units contribute to removal of organic carbon in general and CECs in particular is an open question, especially as the adsorption capacity of the GAC is exhausted. While the AOP step acts to enhance bioavailability of CECs, it is important to assess how the composition of the microbial community influences biodegradative capacity, such as ability to remove CECs, and likewise, how upstream AOPs and water chemistry act to shape the composition of resident microbes and influence this capacity.

Thus far, there are few reported studies devoted to characterizing microbial populations and their activities in large-scale BAC-based BAF systems. However, these characterizations have primarily relied on low resolution measurements and markers for biological activity, such as: oxygen uptake [8], acetate uptake [8,9], culture-based methods [10], total cell counts [9,11,12], ATP [11–13], or no biological analysis at all [14]. Interestingly, total biomass measures do not consistently correlate with BAC treatment performance, in terms of dissolved organic carbon removal efficiency [15,16]. Such observations underscore the need for a deeper, higher resolution understanding of the role of microbial communities within BAF systems [17]. NGS, specifically 16S rRNA gene amplicon sequencing, provides a powerful approach to achieving this by enabling comprehensive profiling of microbial community compositions. 16S rRNA gene amplicon sequencing has previously been applied to characterize advanced treatment processes [18,19]. However, even in the context of drinking water, where BAC-based treatment is much more common and better characterized, available information lacks temporal characterization to map maturation and dynamic community shifts, leaving questions to the extent to which the observed phenomena can be extrapolated to inform potable water reuse system design and operation [8].

Here we utilized 16S rRNA gene amplicon sequencing to characterize the bacterial microbiota characteristic of each stage of treatment in a carbon-based potable reuse train employing coagulation, flocculation, and sedimentation (Floc-Sed), ozonation, BAC filtration, GAC contacting, and UV disinfection. The overarching objective was to demonstrate the potential for microbial community profiling as a diagnostic tool for assessing the functionality and resilience of each treatment stage and to identify specific taxa that are indicative of treatment performance and key water quality characteristics. The specific objectives were to (Table 5.1): 1. establish the extent to which various physical, chemical, and biological processes shape the microbial community composition; 2. identify which bacterial Families are enriched by each stage of treatment; 3. benchmark treatment performance and water quality data to specific bacterial taxa; 4. compare shifts occurring during dynamic operation with baseline temporal shifts in BAC and GAC units. The findings of this study improve understanding of bacterial dynamics occurring in advanced water treatment trains and help to factor this understanding into improved system design and operation.

Table 5.1: Overview of research objectives, hypotheses, and general observations.

Objectives	Hypothesized Results	General Observation(s)	Method(s) of Evaluation
1. Establish the extent to which various physical, chemical, and biological processes shape the microbial community composition through a carbon-based	Chemical treatments, such as ozonation, will induce the strongest shifts, followed by biological and physical treatments, respectively	Microbial community composition is strongly separated by treatment stage. Ozonation resulted in the strongest, and most variable, shift between stages	PCoA Beta Diversity Analysis and ANOSIM

potable reuse treatment train			
2. Identify which bacterial Families are enriched by each stage of physical, chemical, and biological treatment	Each treatment process and variations therein will enrich or attenuate distinct Families of bacteria	Each treatment stage selects for microbes that are indicative of predominant microbiological processes occurring in that stage	PCoA Beta Diversity Analysis and identification of core taxa and comparison of relative abundances through each treatment stage
3. Benchmark treatment performance and water quality data to specific bacterial taxa	Certain taxa serve as effective indicators of treatment performance and resulting water quality	Functions associated with taxa enriched (or not) by altering treatment conditions are consistent with water quality (e.g., lack of ammonia removal in GAC units and lack of enrichment of ammonia-oxidizing bacteria)	Correlation and correspondence analysis of taxa with manipulated variables (e.g., O_3:TOC ratio), incidental variables (e.g., pilot runtime) and water quality measures
4. Compare shifts occurring during dynamic operation to baseline temporal shifts in BAC and GAC units	Ecological succession will occur in microbially-enriched processes in the absence of changing operational conditions	Changes in operational conditions and other disturbance events increase the number and taxa that shift and the magnitude of their relative abundances relative to baseline	Analysis of trends in baseline microbial community composition with time and in response to operational disturbances

METHODS

Site description, sample collection, and preservation

Samples were collected from the SWIFT Pilot facility at the York River Wastewater Treatment Plant (Figure 5.1) over approximately an 18-month period. The SWIFT Pilot treats tertiary (denitrified secondary) effluent and employs: coagulation-flocculation-sedimentation (Floc-Sed), ozonation, BAC filtration, GAC contacting, and UV disinfection. To assess the impact of EBCT, the pilot flow was split between a low EBCT stream of 5 min for the BAC stage (BAC5) and 10 min for the GAC stage (GAC10) or a high EBCT stream of 10 min for the BAC stage (BAC10) and 20 min for the GAC stage (GAC20). In addition to the treatment processes themselves, various chemicals and operational conditions were also adjusted periodically. Additional details on SWIFT pilot operation during this study period are available in prior publications [20–24].

Samples were collected on site and transported on ice to Virginia Tech in sterile 1-L polypropylene bottles. Processing of collected samples occurred within 24 h of collection and included: concentration onto 0.22- µm mixed cellulose esters membrane filter (Millipore, Billerica, MA), fragmentation via sterilized tweezers, and storage at – 20°C. DNA was extracted from filters using a FastDNA SPIN Kit for Soil (MP Biomedicals, Solon, OH) according to the manufacturer's instructions. Extracts were stored at -20 °C (short term) or -80 °C (long term) prior to downstream analysis.

16S rRNA Gene Amplicon Sequencing

Samples were subjected to PCR amplification using the 515f/926r universal bacterial/archaeal primer set targeting the V4-V5 region of the 16S rRNA gene. Following PCR amplification, each sample consisted of pooled triplicate positive PCR products and a negative control, which was used to provide QA/QC validation for sample contamination during PCR amplification. Pooled, positive samples that passed QA/QC were concentrated using the high sensitivity assay (Invitrogen™ Qubit™ 3 Fluorometer, Waltham, MA) before being further pooled with other barcoded samples at 240 ng of DNA per sample. Each final pool contained approximately 150 barcoded samples and were purified using a QIAquick PCR Purification Kit (Qiagen, Valencia, CA) prior to being submitted for sequencing. The Fralin Life Sciences Institute Genomic Sequencing Center at Virginia Tech (Blacksburg, VA) performed amplicon sequencing on an Illumina MiSeq using a 250 cycle paired end protocol. Blanks (including field blanks and trip blanks) were included on each lane of sequencing and, because of low DNA concentrations following PCR amplification, were instead pooled using a max volume protocol.

Reads were processed using the QIIME2 pipeline (Bolyen et al. 2019) and annotated using the Greengenes [26] database (May 2013 release). Libraries were demultiplexed, trimmed to remove low quality reads, and processed on a per lane basis using DADA2 for denoising, chimera removal, and error correction. Amplicon sequence variants generated by DADA2 were merged from each lane of sequencing, clustered at 99%, and classified using the Greengenes database. Classified OTUs were then binned if they contained the same full classification and nontarget sequences (e.g., sequences related to mitochondria and chloroplast) were removed. OTUs were then exported to be utilized for downstream data analysis.

Data Analysis

Beta Diversity, Principal Coordinate Analysis, and Correspondence Analysis

QIIME2's core-metrics and core-metrics phylogenetic plugins were utilized to generate beta-diversity metrics. UniFrac distance matrices (both unweighted and weighted) were analyzed to directly compare the phylogenetic relationships among microbial communities represented by the sequenced samples, which were rarified at both 3,500 and 10,000 reads, to allow for analysis with and without sample that yielded low quantities of reads. Data from the 3,500 rarefaction are presented. Distance matrixes of interest were imported into R version 4.0.2 (Rstudio Team, 2020) to generate principal coordinate analysis (PCoA) plots using the QIIME2r (Bisanz, 2018) and visualized with the ggplot2 packages [28] in R-Studio. Additionally, correspondence analysis was conducted using the vegan and phyloseq packages in R with p-value cutoffs of 0.05 for most comparisons. A more stringent p-value cutoff of 0.01 was applied to higher resolution comparisons (e.g. Family and Genus level) for analysis of BAC and GAC filters.

Identification of Microbial Taxa

Sequences processed in QIIME2 (classified, clustered, condensed, and filtered) were imported into R-v4.0.2 using the QIIME2r package to identify microbial taxa of interest and their changes in abundance both spatially, temporally, and in relation to changes in operational conditions and water quality parameters. Unweighted and weighted analyses were conducted on reads averaged together from triplicate samples. Classified OTUs were analyzed at various classification levels ranging from the phylum to genus level. Family level analysis was selected as the focus of this study, which provided a good balance between phylogenetic resolution and high percent of classified reads (SI Table A.1).

On a weighted basis, abundant taxa were defined as representing at least one percent of all sequenced reads in at least one sample event or condition for each specific treatment process. These taxa were then qualitatively benchmarked to changes in process configuration and upset events and further correlated to extensive water quality and operational data compiled during testing. Further, changes in these abundant taxa during treatment (e.g., changes in the relative abundance of taxa entering and exiting a specific treatment process) were calculated on a per sampling event basis. Taxa that were consistently found to exhibit net increases or decreases in their relative abundance in greater that 66% of sampling events were identified with those experiencing consistent changes in greater than 83% of sampling events corresponding to the relevant analysis. All taxa, not just those meeting the criteria for abundant taxa, were considered in addressing changes in microbial succession/selection during BAC filtration and GAC contacting.

Collection of Water Quality and Operational Data

Abundant taxa identified in the amplicon sequencing data were benchmarked and correlated to operational conditions and water quality data. These measurements utilized a variety of standard methods and were supplemented with continuous, online data for each treatment process via inline sensors for key operational parameters.

Statistical Analysis

All statistical analyses were conducted using the vegan package [29] in R-Studio. Unweighted UniFrac distance matrices were tested using the analysis of similarities (ANOSIM) for categorical variables and adnois test for continuous variables. All statistical testing was

considered significant with a $p < 0.05$. Results of all statistical analyses can be found in SI spreadsheet and Appendix F. Correlation analysis relied on Spearman's rank-order correlation test. Corrections for false discoveries associated with multiple comparisons were accounted for via the Benjamini-Hockberg (BH) procedure for each set of tests. The false discovery rate was liberally set at 0.20 per exploratory recommendations [30], with original p-values reported for those deemed significant after BH correction.

RESULTS

Each Stage of Treatment Uniquely Shapes Associated Microbiota

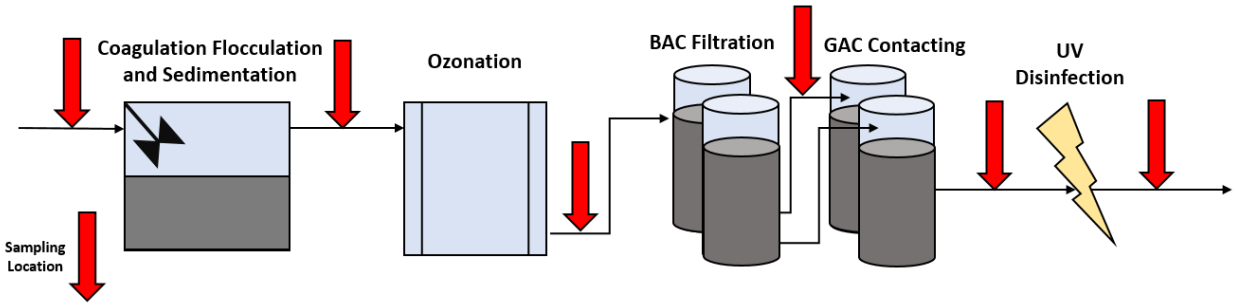


Figure 5.1: HRSD SWIFT Pilot Facility's process schematic with red arrows indicating sampling locations. The pilot influent consisted of effluent from the York River wastewater treatment plant at a flowrate of 4.3 gallons per minute.

Full Treatment Train

PCoA ordination of unweighted and weighted UniFrac distances for the entire pilot treatment train are presented in Figure 5.2a and Figure 5.2b, respectively. Overall, unweighted and weighted analyses indicated highly similar trends. Therefore, subsequent beta diversity analysis reported here focus on unweighted analysis, unless noted otherwise. When analyzing the entire dataset, all factors tested were found to be significantly associated with the composition of the bacterial microbiota; however, variation in the bacterial microbiota was best explained (i.e., highest r-statistic or R2 value) by treatment process (ANOSIM $p < 0.05$; r-stat = 0.52). Sampling event (ANOSIM $p < 0.05$; r-stat = 0.23), temperature (Adonis $p < 0.05$; R2 = 0.04), and pilot run time (Adonis $p < 0.05$; R2 = 0.03) were also found to be significant factors, but accounted for a much lower degree of separation among the microbial communities.

From the PCoA analysis, clustering was apparent between samples from the pilot influent and Floc-Sed effluent, with additional clustering of GAC10 and GAC20 effluents, indicating large-scale shifts in microbial composition through the treatment train (Figure 5.2). Notably, the bacterial microbiota associated with ozonation and BAC filtration (both BAC5 and BAC10) displayed the highest levels of variability and accounted for the sharpest transitions in microbial community composition. UV effluent displayed a two-part clustering, where one group was clustered similarly to the GAC effluent and the other clustered with a subset of BAC effluent samples that were divergent from other major clusters. In the weighted analysis, UV samples did not cluster in two groups nor with GAC effluent, instead overlapping with samples from both the ozonation and BAC treatment processes.

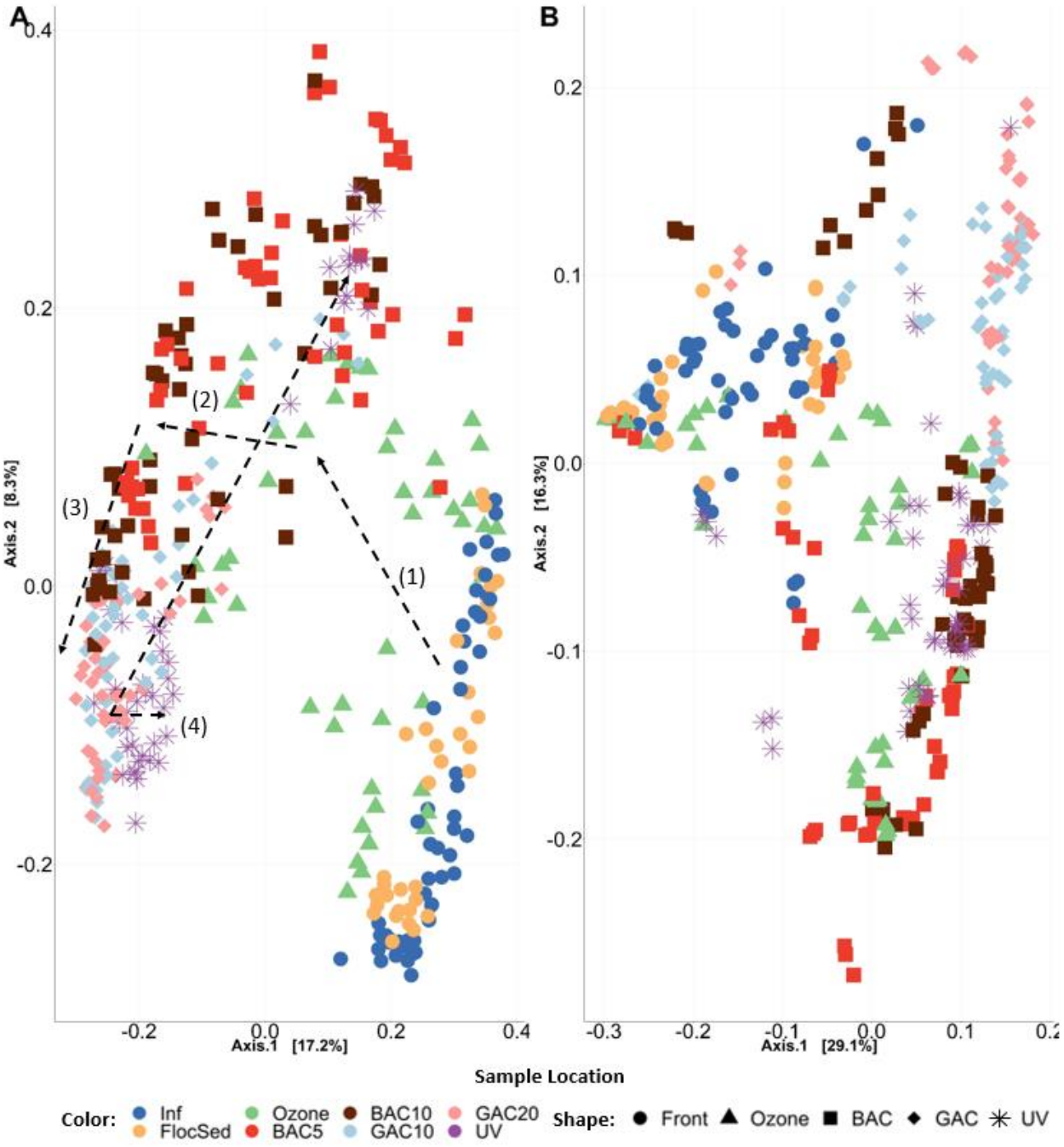


Figure 5.2: UniFrac beta diversity plot comparing effluent samples from each indicated stage of treatment across the AWT pilot based on 16S rRNA gene amplicon sequencing. Analysis includes all sampling points collected with time from each point over 18 months of operation. Samples are classified by their (A) unweighted and (B) weighted based analysis. Dashed arrows indicate the spatial series of sampling locations through the pilot.

Pilot Influent, and Coagulation Flocculation and Sedimentation

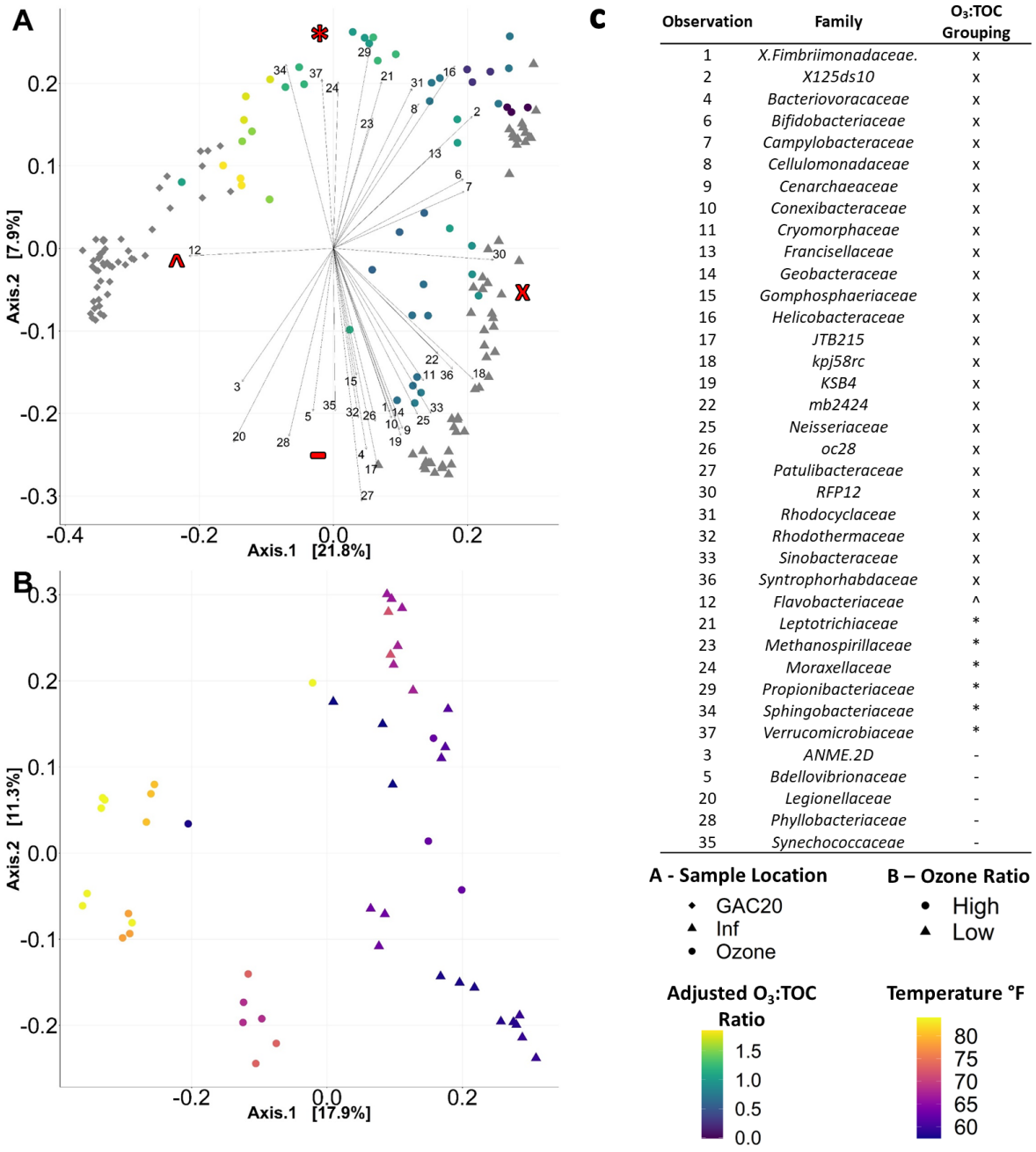
Pilot influent and Floc-Sed effluent were found to have a high level of compositional overlap and were therefore analyzed together. Variation associated with these processes were

best explained by their individual sampling event (ANOSIM $p < 0.05$; r-stat = 0.93) and temperature (Adonis $p < 0.05$; R2 = 0.13) (SI Table F.5-F.6 and Appendix F: Subsection 1). The composition of the bacterial microbiota was found to have poorer associations with treatment process (ANOSIM $p < 0.05$; r-stat = 0.09) or pilot run time (Adonis $p < 0.05$; R2 = 0.057).

Ozonation

The variation in the composition of the bacterial microbiota following ozonation was best explained by sample event (ANOSIM $p < 0.05$; r-stat = 0.99) and then to a lesser extent by ozone:TOC, adjusted for nitrite concentration (Adonis $p < 0.05$; R2 = 0.14) and temperature (Adonis $p < 0.05$; R2 = 0.14) (Figure 5.3ab). Pilot run time, though statistically significant, explained a much lower percentage of compositional variation (Adonis $p < 0.05$; R2 = 0.05). From Figure 5.3ab it can be seen that lower O3:TOC ratios adjusted for nitrite concentrations (and ozone doses) resulted in bacterial microbiota compositions more similar to pilot influent (and Floc-Sed effluent) compositions while higher ratios lead to greater shifts in the microbial composition. O3:TOC = 1 was identified as a threshold distinguishing these two clusters (ANOSIM $p < 0.05$; r-stat = 0.53), Appendix F: Subsection 2.

Correspondence analysis on the unweighted Unifrac analysis at the family level identified associated taxa with O_3:TOC ratios adjusted for nitrite concentration (p-value < 0.05). Specifically, the families Sphingobacteriaceae, Verrucomicrobiacease, Methanosarcinaceae, Leptotrichiaceae, and others were associated with O3:TOC ratios ~1.25 (Figure 5.3ab). Further, adjusted O_3:TOC ratios ~1.00 were associated with Francisellaceae, Bifidobacteriacease, Campylobacteraceae, and others, while ratios ~0.75 were associated with Cryomorphacease, Sinobacteracease, Neisseriaceae, Cenarchaeaceae, and others. Interestingly, no Families were closely associated with adjusted O_3:TOC >1.5 when utilizing an unweighted analysis. The family Flavobacteriaceae was associated with the GAC20 effluent which was provided as a reference community composition indicative of BAF treatment.



Observation	Family	O_3:TOC Grouping
1	*X.Fimbriimonadaceae.*	x
2	*X125ds10*	x
4	*Bacteriovoracaceae*	x
6	*Bifidobacteriaceae*	x
7	*Campylobacteraceae*	x
8	*Cellulomonadaceae*	x
9	*Cenarchaeaceae*	x
10	*Conexibacteraceae*	x
11	*Cryomorphaceae*	x
13	*Francisellaceae*	x
14	*Geobacteraceae*	x
15	*Gomphosphaeriaceae*	x
16	*Helicobacteraceae*	x
17	*JTB215*	x
18	*kpj58rc*	x
19	*KSB4*	x
22	*mb2424*	x
25	*Neisseriaceae*	x
26	*oc28*	x
27	*Patulibacteraceae*	x
30	*RFP12*	x
31	*Rhodocyclaceae*	x
32	*Rhodothermaceae*	x
33	*Sinobacteraceae*	x
36	*Syntrophorhabdaceae*	x
12	*Flavobacteriaceae*	^
21	*Leptotrichiaceae*	*
23	*Methanospirillaceae*	*
24	*Moraxellaceae*	*
29	*Propionibacteriaceae*	*
34	*Sphingobacteriaceae*	*
37	*Verrucomicrobiaceae*	*
3	*ANME.2D*	-
5	*Bdellovibrionaceae*	-
20	*Legionellaceae*	-
28	*Phyllobacteriaceae*	-
35	*Synechococcaceae*	-

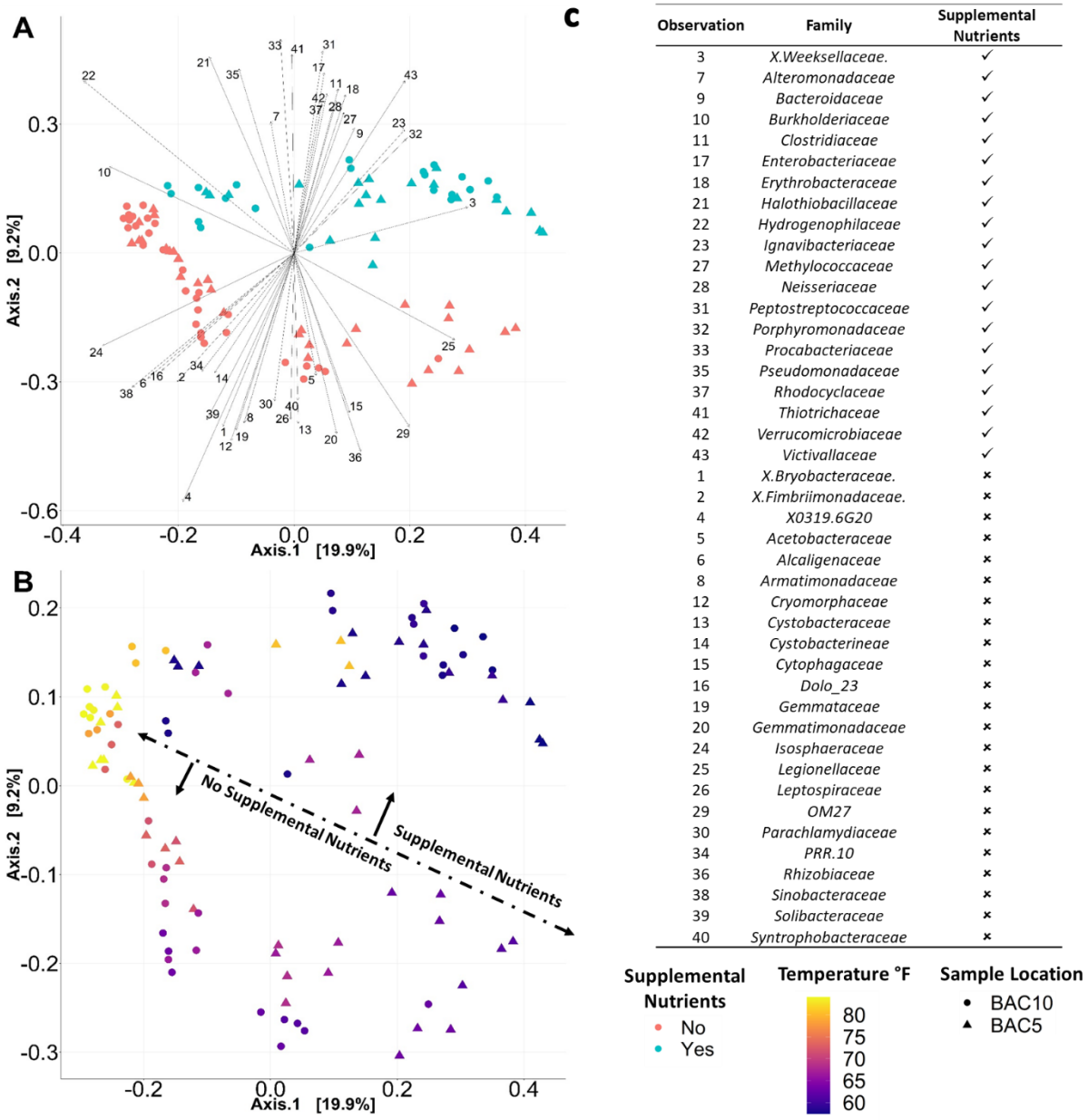
Figure 5.3: Unweighted UniFrac beta diversity plot of ozone effluent microbial communities based on 16S rRNA gene amplicon sequencing. Samples are classified by their (A) O_3:TOC ratio adjusted for nitrite concentration gradient with Influent and GAC20 samples included for comparison and correspondence analysis (B) temperature gradient with categorical classification of ozone ratio (High > 1.0 > Low). (C) Key for the correspondence analysis taxa indicated in (A) and corresponding O_3:TOC groupings identified on the figure (^, *, x, and -).

BAC Filtration

Variation in microbial composition of the BAC filter effluents was best explained by the specific sampling event (BAC5 ANOSIM $p < 0.05$; r-stat = 0.96| BAC10 ANOSIM $p < 0.05$; r-stat = 0.95) and the impact of supplemental nutrient additions in the form of nitrogen and phosphorous (BAC5 ANOSIM $p < 0.05$; r-stat = 0.32 | BAC10 ANOSIM $p < 0.05$; r-stat = 0.54) (Figure 5.4a). Regardless of the impact of nutrient additions, chlorine quenching type prior to BAC filtration (BAC5 ANOSIM $p < 0.05$; r-stat = 0.32 | BAC10 ANOSIM $p < 0.05$; r-stat = 0.36) and the inclusion of polymer to enhance filtration (BAC5 ANOSIM $p < 0.05$; r-stat = 0.25 | BAC10 ANOSIM $p < 0.05$; r-stat = 0.29) were also found to be significant cofactors in explaining variation in bacterial composition (Appendix F: Subsection 3). Additionally, temperature fluctuations experienced during testing was found to account for approximately 15% of the variation in the BAC5 (Adonis $p < 0.05$; $R2 = 0.15$) and approximately 17% of the variation in the BAC10 (Adonis $p < 0.05$; $R2 = 0.17$) effluent bacterial communities (Figure 5.4b). Pilot run time accounted for slightly less variation than temperature (BAC5 Adonis $p < 0.05$; $R2 = 0.07$ | BAC10 Adonis $p < 0.05$; $R2 = 0.13$) while O_3:TOC ratios adjusted for nitrite concentrations were also found to account for similar levels of variation (BAC5 Adonis $p < 0.05$; $R2 = 0.12$ and 0.11 | BAC10 Adonis $p < 0.05$; $R2 = 0.15$ and 0.26) as filter runtime (Appendix F: Subsection 3). Proximity of sampling to backwash (e.g., time since most recent backwash) explained the least variation for both BAC5 (Adonis $p < 0.05$; $R2 = 0.06$) and BAC10 (Adonis $p < 0.05$; $R2 = 0.06$) (Appendix F: Subsection 3).

Given the apparent impact of nutrient additions on the bacterial microbiota (Figure 5.4ab), a subset of samples were analyzed before and after the addition of nutrients. Within this subset, higher R2 values were noted as a function of temperature before (BAC5 Adonis $p < 0.05$; $R2 = 0.28$ | BAC10 Adonis $p < 0.05$; $R2 = 0.25$) than after (BAC5 Adonis $p < 0.05$; $R2 = 0.11$ | BAC10 Adonis $p < 0.05$; $R2 = 0.17$) addition of nutrients. However, temperature was largely controlled following nutrient addition, which limited the latter comparison. Chlorine quenching type similarly affected bacterial community composition before nutrient addition (BAC5 ANOSIM $p < 0.05$; r-stat = 0.12 | BAC10 ANOSIM $p < 0.05$; r-stat = 0.15), but not after. Weighted analysis of chlorine quenching; however, did indicate an effect after nutrient dosing (BAC5 ANOSIM $p < 0.05$; r-stat = 0.45 | BAC10 ANOSIM $p < 0.05$; r-stat = 0.33) indicating that there was an effect on relative abundance. Also similarly, O_3:TOC ratios adjusted for nitrite concentrations, were associated with elevated R2 values for both BAC5 and BAC10 when analyzed after nutrient supplementation (BAC5 Adonis $p < 0.05$; $R2 = 0.13$ | BAC10 Adonis $p < 0.05$; $R2 = 0.20$) and before (BAC5 Adonis $p < 0.05$; $R2 = 0.19$ | BAC10 Adonis $p < 0.05$; $R2 = 0.18$).

Correspondence analysis of the unweighted Unifrac distance matrix highlighted Families that were associated with before versus after addition of supplemental nitrogen/phosphorus to the BAC5 and BAC10 filters (Figure 5.4ab). Specifically, Hydrogenophilaceae, Halothiobacillaceae, Pseudomonadaceae, Verrucomibrobiacease, Enterobacteriaceae, Methylococcaceae, Bacteroidaceae, and Weeksellaceae (among others) were associated with the supplementation of nitrogen and phosphorous. On the other hand, Isophaeraceae, Sinobacteraceae, Solibacteraceae, Armatimonadaceae, Leptospiraceae, Rhosobacteraceae, and Legionellaceae (among others) were found to be prominent in the BAC5 and BAC10 filter effluent prior to nutrient supplementation. Burkholderiaceae was found to be associated with samples collected at the highest recorded temperature, regardless of the addition of nutrients.



Observation	Family	Supplemental Nutrients
3	*X.Weeksellaceae.*	✓
7	*Alteromonadaceae*	✓
9	*Bacteroidaceae*	✓
10	*Burkholderiaceae*	✓
11	*Clostridiaceae*	✓
17	*Enterobacteriaceae*	✓
18	*Erythrobacteraceae*	✓
21	*Halothiobacillaceae*	✓
22	*Hydrogenophilaceae*	✓
23	*Ignavibacteriaceae*	✓
27	*Methylococcaceae*	✓
28	*Neisseriaceae*	✓
31	*Peptostreptococcaceae*	✓
32	*Porphyromonadaceae*	✓
33	*Procabacteriaceae*	✓
35	*Pseudomonadaceae*	✓
37	*Rhodocyclaceae*	✓
41	*Thiotrichaceae*	✓
42	*Verrucomicrobiaceae*	✓
43	*Victivallaceae*	✓
1	*X.Bryobacteraceae.*	×
2	*X.Fimbriimonadaceae.*	×
4	*X0319.6G20*	×
5	*Acetobacteraceae*	×
6	*Alcaligenaceae*	×
8	*Armatimonadaceae*	×
12	*Cryomorphaceae*	×
13	*Cystobacteraceae*	×
14	*Cystobacterineae*	×
15	*Cytophagaceae*	×
16	*Dolo_23*	×
19	*Gemmataceae*	×
20	*Gemmatimonadaceae*	×
24	*Isosphaeraceae*	×
25	*Legionellaceae*	×
26	*Leptospiraceae*	×
29	*OM27*	×
30	*Parachlamydiaceae*	×
34	*PRR.10*	×
36	*Rhizobiaceae*	×
38	*Sinobacteraceae*	×
39	*Solibacteraceae*	×
40	*Syntrophobacteraceae*	×

Figure 5.4: Unweighted UniFrac beta diversity plot for both BAC5 and BAC10 effluent samples with sample event identified. Samples are classified by their (A) relation to supplemental nutrient additions and correspondence analysis (B) temperature gradient. (C) Provides a reference table for the correspondence analysis grouping identified taxa with supplemental nutrient addition.

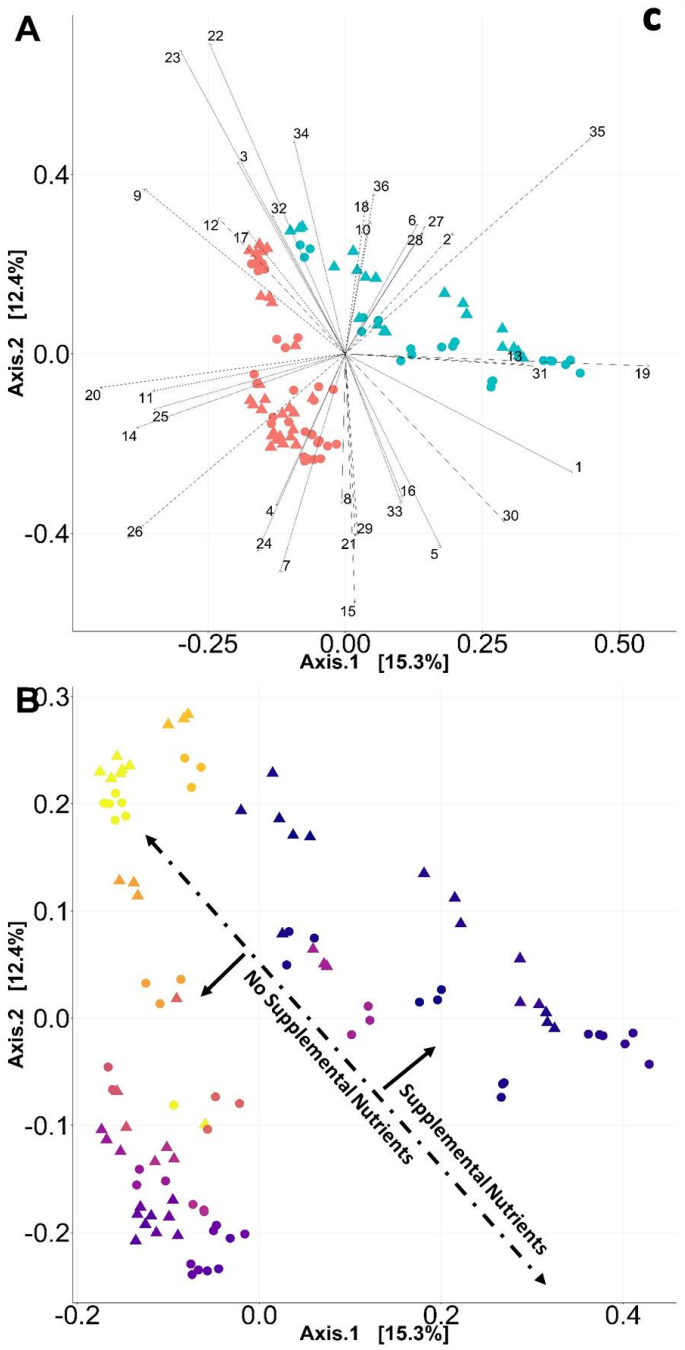
GAC Contacting

Similarly to the BAC filters, succession of the bacterial microbiota associated with GAC contacting was most strongly related to changes in operational conditions and water quality parameters, like temperature (Figure 5.5ab). Specifically, the sampling event (GAC10 ANOSIM $p < 0.05$; R-stat = 0.98 | GAC20 ANOSIM $p < 0.05$; R-stat = 0.99), supplemental nutrients in the

form of nitrogen and phosphorus (GAC10 ANOSIM p <0.05; R-stat = 0.54 | GAC20 ANOSIM p <0.05; R-stat = 0.63), and the inclusion of filter aid polymer (GAC10 ANOSIM p <0.05; R-stat = 0.53 | GAC20 ANOSIM p <0.05; R-stat = 0.57) were found to account for the highest levels of variation. Unlike BAC filtration, the type of chlorine quenching agent fed prior to BAC filtration (GAC10 ANOSIM p <0.05; R-stat = 0.26 | GAC20 ANOSIM p <0.05; R-stat = 0.29) did not have as strong as an effect as the other factors. When considering continuous variables, temperature (Adonis GAC10 Adonis p <0.05; R2 = 0.14 | GAC20 Adonis p <0.05; R2 = 0.12), pilot run time (Adonis GAC10 Adonis p <0.05; R2 = 0.15 | GAC20 Adonis p <0.05; R2 = 0.16), and applied O_3:TOC ratios adjusted for nitrite concentrations (Adonis GAC10 Adonis p <0.05; R2 = 0.14 | GAC20 Adonis p <0.05; R2 = 0.13) were found to be the most strongly related to bacterial community composition. All other variables accounted for roughly 15% of the microbial variation. Conversely, the temporal proximity to the most recent backwash event was found to be less effective at explaining microbial variation (Adonis GAC10 Adonis p <0.05; R2 = 0.06 | GAC20 Adonis p <0.05; R2 = 0.08) (Appendix F: Subsection 4).

Additional trends emerged when the data were analyzed as subsets before and after nutrient addition. A stronger effect of the type of chlorine quenching agent fed to the pilot was noted when GAC data were analyzed as a subset of samples collected after the addition of nutrients (GAC10 ANOSIM p <0.05; R-stat = 0.32 | GAC20 ANOSIM p <0.05; R-stat = 0.76). Temperature, when subsetted for samples prior to the additions of supplemental nutrients, was also found to account for higher levels of GAC effluent community variation (Adonis GAC10 Adonis p <0.05; R2 = 0.25 | GAC20 Adonis p <0.05; R2 = 0.25). Temperatures recorded prior to supplementing nutrients were also more variable (temperature control began roughly 1 month after nutrient supplementation). Additionally, the impact of variable nitrite adjusted O3:TOC ratios accounted for increased community variability when subsetted before (Adonis GAC10 Adonis p <0.05; R2 = 0.21 | GAC20 Adonis p <0.05; R2 = 0.23) and after (Adonis GAC10 Adonis p <0.05; R2 = 0.21 | GAC20 Adonis p <0.05; R2 = 0.19) nutrient addition with the former generally subjected to higher O3:TOC ratios.

Correspondence analysis on the unweighted Unifrac distance matrix before and after nutrient addition identified Families associated with each factor (Figure 5.5ab). Alcaligenaceae, Leptospiraceae, Bacillaceae, Simkaniaceae, Conexibacteraceace, Nisseriaceae, Chthonomonadaceae (among others) were associated with GAC effluent before the addition of nutrients. Rhodocyclaceae, Erythrobacteraceae, Clostridiaceae, Verrucomicrobiaceae, Peptosteptococcaceae, Procabacteriaceae, Enterobacteriaceae, Thiotrichaceae, Halothiobacillaceae, Hydrogenophilaceae, Aeromonadaceae, and Victivallaceae (among others) appeared to be enriched by the supplementationof nutrients.



Observation	Family	Supplemental Nutrients
2	*X.Chromatiaceae.*	✓
3	*X.Fimbriimonadaceae.*	✓
6	*Aeromonadaceae*	✓
10	*Bacteroidaceae*	✓
13	*Clostridiaceae*	✓
18	*Enterobacteriaceae*	✓
19	*Erythrobacteraceae*	✓
22	*Halothiobacillaceae*	✓
23	*Hydrogenophilaceae*	✓
27	*Peptostreptococcaceae*	✓
28	*Procabacteriaceae*	✓
31	*Rhodocyclaceae*	✓
32	*Rhodothermaceae*	✓
34	*Thiotrichaceae*	✓
35	*Verrucomicrobiaceae*	✓
36	*Victivallaceae*	✓
1	*X.Cerasicoccaceae.*	-
5	*X0319.6G20*	-
16	*Cystobacterineae*	-
30	*Rhodobacteraceae*	-
33	*Syntrophobacteraceae*	-
4	*X.Weeksellaceae.*	×
7	*Alcaligenaceae*	×
8	*Bacillaceae*	×
9	*Bacteriovoracaceae*	×
11	*Cenarchaeaceae*	×
12	*Chthonomonadaceae*	×
14	*Conexibacteraceae*	×
15	*Cystobacteraceae*	×
17	*Ectothiorhodospiraceae*	×
20	*Gaiellaceae*	×
21	*Gemmataceae*	×
24	*Leptospiraceae*	×
25	*Neisseriaceae*	×
26	*OM27*	×
29	*Rhizobiaceae*	×

Figure 5.5: Unweighted UniFrac beta diversity plot for both GAC10 and GAC20 effluent samples with sample event identified. Samples are classified by their (A) relation to supplemental nutrient additions and correspondence analysis (B) temperature gradient. (C) Provides a reference table for the correspondence analysis grouping identified taxa with supplemental nutrient addition.

UV Disinfection

Other than sampling event (ANOSIM $p < 0.05$; r-stat = 0.95) (Appendix F: Subsection 5), chlorine quenching was found to be the strongest factor shaping the microbiota of the UV

effluent (ANOSIM $p < 0.05$; r-stat = 0.42). Nutrient addition was the next most impactful factor (ANOSIM $p < 0.05$; r-stat = 0.19), followed by polymer addition (ANOSIM $p < 0.05$; r-stat = 0.10). Of the other tested variables, none explained >12% of the variability, with most comparisons resulting in $R2 < 0.10$. Pilot run time was identified as the strongest factor (Adonis $p < 0.05$; $R2 = 0.12$).

Assessment of Process Conditions, Operational data, and Water Quality Parameters on Taxa Abundances

Abundant families consistently selected for or removed during each treatment process are reported in SI datasheet. Table 5.2 presents the specific subset of these data corresponding to the number of abundant taxa that exhibited consistent changes in their relative abundance in >66% of sampling events with taxa experiencing consistent changes in >83% of sampling events identified by name and discussed in later sections. In addition to taxa exhibiting consistent changes in abundance over the full testing period, similar taxa were also identified when subsetting the data before and after the addition of nutrients (Table 5.3).

Flocculation and Sedimentation

A full correlation analysis between core taxa and operational data can be found in SI datasheet. Briefly, effluent turbidity was identified as being most strongly correlated to relative abundances of dominant families found within Floc-Sed effluent. Turbidity was also found to be the strongest indicator of treatment performance. Effluent turbidity was significantly correlated with 6 core families, all of which were also noted above as either consistently increasing or decreasing in greater than 66% of sampling events. Namely Flavobacteriaceae (Spearman p-value < 0.05; R2 = 0.79), kpj58rc (Spearman p-value < 0.05; R2 = -0.78), Pseudomonadaceae (Spearman p-value < 0.05; R2 = -0.71), Rhizobiaceae (Spearman p-value < 0.05; R2 = -0.89), Bacteriovoracaceae (Spearman p-value < 0.05; R2 = -0.76), and Bdellovibrionaceae (Spearman p-value < 0.05; R2 = -0.77). Additionally, the taxa of interest Nitrospireace was found to be one of two core families well correlated with effluent TOC (Spearman p-value < 0.05; R2 = -0.85), the other being Hydrogenophilaceae (Spearman p-value < 0.05; R2 = -0.72).

Table 5.2: Identified Families that underwent consistent selection or removal during treatment.

Treatment Process	# of Families Reduced in > 66% Sampling Events	Familes Consistently Removed in > 83% of Sampling Events	# of Families Enriched in > 66% Sampling Events	Families Consistently Enriched in > 83% of Sampling Events
Floc-Sed*	11	Saprospiraceae, Flavobacteriaceae, Cryomorphaceae, Haliangiaceae, Rhodocyclaceae, Xanthomonadaceae, Polyangiaceae, and Methylophilaceae	14	Legionellaceae, Oxalobacteraceae, Bdellovibrionaceae, Bacteriovoracaceae, Neisseriaceae, Rhizobiaceae, Sphingobacteriaceae, and Procabacteriaceae
Ozone*	15	Thiotrichaceae, Ruminococcaceae, Procabacteriaceae, Methylophilaceae, Moraxellaceae, Sphingobacteriaceae, Coxiellaceae, kpj58rc, Neisseriaceae, Alteromonadaceae, and Rhodocyclaceae	19	Hyphomonadaceae, Oxalobacteraceae, Caulobacteraceae, Nitrosomonadaceae, Rhodospirillaceae, Sphingomonadaceae, Bradyrhizobiaceae, Cytophagaceae, Hyphomicrobiaceae, Acetobacteraceae, Planctomycetaceae, Mycobacteriaceae, and Methylobacteriaceae
BAC5	24	Pseudomonadaceae, and Xanthomonadaceae	10	Oxalobacteraceae and Rhabdochlamydiaceae
BAC10	20	Clostridiaceae, and Sphingobacteriaceae	19	Bdellovibrionaceae, Coxiellaceae, Cystobacterineae, oc28, Opitutaceae, Phyllobacteriaceae, and Pirellulaceae
Both BAC5 and BAC10	-	Methylobacteriaceae, Methylophilaceae, Mycobacteriaceae, Nannocystaceae, Peptostreptococcaceae, Procabacteriaceae, Rhodocyclaceae, and Saprospiraceae	-	Hyphomicrobiaceae
GAC10	8	Sphingomonadaceae, and Comamonadaceae	25	Bdellovibrionaceae, Parachlamydiaceae, Phyllobacteriaceae and Pirellulaceae
GAC20	12	Bradyrhizobiaceae, Caulobacteraceae, Flavobacteriaceae, Nitrosomonadaceae, and oc28	26	Chitinophagaceae, Gemmataceae, Legionellaceae, Rhabdochlamydiaceae, and Syntrophobacteraceae
Both GAC10 and GAC20	-	Oxalobacteraceae	-	0319-6G20, Acetobacteraceae, Caldilineaceae, Coxiellaceae, Cryomorphaceae, Dolo_23, Nitrospiraceae, Opitutaceae, Planctomycetaceae, PRR-10, R4-41B, Rhodospirillaceae, Rickettsiaceae, and Sphingobacteriaceae
UV	29	Opitutaceae, Planctomycetaceae, Pirellulaceae, Rhodospirillaceae, Bdellovibrionaceae, Coxiellaceae, Bryobacteraceae, Rhabdochlamydiaceae, Phyllobacteriaceae, Rickettsiaceae, Cryomorphaceae, Caldilineaceae, Hyphomicrobiaceae, Legionellaceae, Nitrospiraceae, Nitrosomonadaceae, and Procabacteriaceae	8	Sphingomonadaceae, Mycobacteriaceae, Comamonadaceae, and Rhodocyclaceae

*Due to decreased number of sampling events the threshold for consistant selection was set at 75% instead of 83%.

Ozonation and UV Disinfection

Correlation analysis revealed 12 abundance taxa significantly correlated with average ozone dose and 11 with O3:TOC ratio adjusted for nitrite concentrations with 10 taxa shared between both operational conditions. This included the families Bdellovibrionaceae

(Spearman p-value < 0.05; ozone dose R2 = 0.53 and O3:TOC ratio R2 = 0.58), Bradyrhizobiaceae (Spearman p-value < 0.05; ozone dose R2 = 0.66 and O3:TOC ratio R2 = 0.76), Comamonadaceae (Spearman p-value < 0.05; ozone dose R2 = 0.71 and O3:TOC ratio R2 = 0.73), Cytophagaceae (Spearman p-value < 0.05; ozone dose R2 = 0.81 and O3:TOC ratio R2 = 0.78), Flavobacteriaceae (Spearman p-value < 0.05; ozone dose R2 = -0.52 and O3:TOC ratio R2 = -0.60), Methylobacteriaceae (Spearman p-value < 0.05; ozone dose R2 = 0.81 and O3:TOC ratio R2 = 0.63), Methylophilaceae (Spearman p-value < 0.05; ozone dose R2 = -0.61 and O3:TOC ratio R2 = -0.81), Mycobacteriaceae (Spearman p-value < 0.05; ozone dose R2 = 0.51 and O3:TOC ratio R2 = 0.45), Nannocystaceae (Spearman p-value < 0.05; ozone dose R2 = -0.55 and O3:TOC ratio R2 = -0.67),Nocardiaceae (Spearman p-value < 0.05; ozone dose R2 = 0.71 and O3:TOC ratio R2 = 0.71), Peptostreptococcaceae (Spearman p-value < 0.05; ozone dose R2 = -0.48 and O3:TOC ratio R2 = -0.66), and Procabacteriaceae (Spearman p-value < 0.05; ozone dose R2 = -0.63 and O3:TOC ratio R2 = -0.67) of which Bdellovibrionaceae, Methylophilaceae, and Procabacteriaceae were found to be reduced in greater than 66% of sampling events. Additionally, UVT, temperature, and bromate concentrations were also found to be relatively well correlated with abundant taxa's relative abundance identifying 10, 12, and 13 taxa, respectively. Taxa that correlated with UVT included Bdellovibrionaceae (Spearman p-value < 0.05; R2 = 0.68), Methylophilaceae (Spearman p-value < 0.05; R2 = -0.83), Neisseriaceae (Spearman p-value < 0.05; R2 = -0.71), and Procabacteriaceae (Spearman p-value < 0.05; R2 = -0.76), all of which decreased in relative abundance reductions in >66% of sampling events. Bradyrhizobiaceae (Spearman p-value < 0.05; R2 = 0.60), Comamonadaceae (Spearman p-value < 0.05; R2 = 0.61), Cytophagaceae (Spearman p-value < 0.05; R2 = 0.56), Oxalobacteraceae (Spearman p-value < 0.05; R2 = 0.80), and Peptostreptococcaceae (Spearman p-value < 0.05; R2 = -0.62) also correlated with UVT and consistently increased in relative abundance. Comamonadaceae also strongly correlated with temperature (Spearman p-value < 0.05; R2 = 0.75) and bromate concentrations (Spearman p-value < 0.05; R2 = 0.68).

There was largely a lack of correlation between abundant taxa and operational conditions or water quality parameters. UV254 failed to correlate significantly with any abundant taxa while UVT was only correlated with two families Peptostreptococcaceae (Spearman p-value < 0.05; R2 = -0.62) and Verrucomicrobiaceae (Spearman p-value < 0.05; R2 = 0.65), neither of which exhibited consistent changes over multiple sampling events. Legionellaceae and Mycobacteriaceae were also not found to be significantly correlated with any tested parameter. When considering other taxa that were consistently enriched in the pilot, there were no identifiable correlations with Sphingomonadaceae or Rhodocyclaceae. On the other hand, Comamonadaceae was found to be significantly correlated with alkalinity (Spearman p-value < 0.05; R2 = -0.71), calcium concentrations (Spearman p-value < 0.05; R2 = -0.63), and hardness (Spearman p-value < 0.05; R2 = -0.73), among others.

BAC Filtration

When correlated to water quality parameters and operational conditions, effluent nitrogen concentrations (specifically NH4, NO2, and NO3), TOC removal percentage, and adjusted O3:TOC ratio were notable. The most taxa were found to be correlated with adjusted O3:TOC ratio (14 and 16) and NH4 effluent concentration (13 and 10) for BAC5 and BAC10, respectively. Few taxa were correlated with NO2 (6 and 3), NO3 (3 and 1) or TOC removal percentage (4 and 5) for BAC5 and BAC10, respectively. For BAC5, TOC removal percentage

was correlated with the families Bacteriovoracaceae (Spearman p-value < 0.05; R2 = 0.71), Coxiellaceae (Spearman p-value < 0.05; R2 = 0.53), Oxalobacteraceae (Spearman p-value < 0.05; R2 = -0.59), and Pseudomonadaceae (Spearman p-value < 0.05; R2 = -0.66), all of which were identified as significantly shifting in abundance in 66% or more of sampling events, while BAC10 was correlated with Clostridiaceae (Spearman p-value < 0.05; R2 = -0.71), Oxalobacteraceae (Spearman p-value < 0.05; R2 = -0.58), Pirellulaceae (Spearman p-value < 0.05; R2 = -0.77), Pseudomonadaceae (Spearman p-value < 0.05; R2 = -0.67), and Sphingomonadaceae (Spearman p-value < 0.05; R2 = -0.74) where 3 of the 5 were identified previously. Interestingly, Nitrosomonadaceae was found to only be correlated with effluent NO3 concentrations (Spearman p-value < 0.05; R2 = 0.56) in BAC5 and was not significantly correlated with any parameters in BAC10. Similarly, Nitrospiraceae was not correlated with any test parameters in BAC5 (it did not meet the abundance threshold) and only correlated with pilot run time (Spearman p-value < 0.05; R2 = 0.67) in BAC10.

Table 5.3: Identified Families that underwent increased selection or removal during BAC and GAC treatment following supplementation nutrient additions.

Treatment Process	Increased Enrichment Following Nutrient Additions	Increased Enrichment Prior to Nutrient Additions
Both BAC5 and BAC10	Bacteriovoracaceae, Flavobacteriaceae, Halothiobacillaceae, Nitrospiraceae, Pseudomonadaceae, Sphingobacteriaceae, Sphingomonadaceae, and Hydrogenophilaceae	Erythrobacteraceae, Rhizobiaceae, Rhodocyclaceae, Bdellovibrionaceae, Cytophagaceae, and Rhodobacteraceae
Both GAC10 and GAC20	A4b, Bacteriovoracaceae, Cytophagaceae, and Hydrogenophilaceae	Cerasicoccaceae, Hyphomonadaceae, and Procabacteriaceae

GAC Contacting

Correlation analysis revealed several operational conditions and water quality parameters that were significantly correlated for both GAC10 and GAC20. Specifically, after GAC10, effluent TOC (13), filter runtime (13), temporal proximity to backwash (12), effluent NH4 concentration (12), pilot runtime (10), adjusted O3:TOC ratio (10), change in NH4 concentration (8), and temperature (8) were found to be most strongly correlated with multiple, abundant bacterial families. At GAC20, similar relationships were found between pilot runtime (10), temperature (9), adjusted O3:TOC ratio (9), effluent TOC (8), temporal proximity to backwash (8), filter runtime (8), and effluent NH4 concentration (7) and dominant families. A closer look at taxa identified above, specifically the 83% sampling event threshold, revealed consistent correlations in both GAC10 and GAC20 effluents. Oxalobacteraceae, which experienced consistent relative abundance reductions, was negatively correlated with temperature for both GAC10 (Spearman p-value < 0.05; R2 = -0.93) and GAC20 (Spearman p-value < 0.05; R2 = -0.83). Acetobacteracae was found to be negatively correlated with effluent NO2 concentrations for both GAC10 (Spearman p-value < 0.05; R2 = -0.60) and GAC20 (Spearman p-value < 0.05; R2 = -0.72) with additional relationships with NO3 (Spearman p-value < 0.05; R2 = -0.70) and TIN concentrations (Spearman p-value < 0.05; R2 = -0.75) only at GAC20 and wide range of additional parameters at GAC10 including NH4 concentrations (Spearman p-value < 0.05; R2 = 0.63), effluent TOC (Spearman p-value < 0.05; R2 = -0.83), and filter runtime (Spearman p-value < 0.05; R2 = 0.65), among others. Caldilineacae was found to be negatively correlated with

pilot runtime for both GAC10 (Spearman p-value < 0.05; R2 = -0.90) and GAC20 (Spearman p-value < 0.05; R2 = -0.84) with additional correlates between effluent TOC (Spearman p-value < 0.05; R2 = -0.73) at GAC10 and effluent NH_4 concentrations (Spearman p-value < 0.05; R2 = 0.82) at GAC20. Cryomorphacae was found to be correlated with adjusted O3:TOC ratios (Spearman p-value < 0.05; R2 = 0.66, R2 = 0.63) and effluent TOC concentrations (Spearman p-value < 0.05; R2 = -0.66, R2 = -0.76) for both GAC10 and GAC20, respectively. Rhodospirillacae was also found to have consistent correlations at GAC10 and GAC20, specifically with filter runtime (Spearman p-value < 0.05; R2 = 0.67, R2 = 0.56), pilot runtime (Spearman p-value < 0.05; R2 = -0.81, R2 = -0.66), temporal proximity to backwash (Spearman p-value < 0.05; R2 = 0.59, R2 = 0.58), and effluent TOC (Spearman p-value < 0.05; R2 = -0.73, R2 = -0.70). Rickettsiaceae and Sphingobacteriacae were also found to experience correlated with nitrogen concentrations only at GAC10 with both taxa being correlated with effluent NH_4 concentrations (Spearman p-value < 0.05; R2 = 0.83, R2 = 0.71, respectively) and Sphingobacteriacae additionally being correlated with effluent NO_3 concentrations (Spearman p-value < 0.05; R2 = 0.67). Interestingly, the taxa of interest Nitrospiracae, Nitrosomonadaceae, and Planctomycetaceae were not correlated with any tested parameters at either GAC10 or GAC20.

Identification of Temporal Succession and Steady State BAF Operation

The cumulative, temporal changes in the relative abundance of phylum, family, and genus classified taxa can be found in Figure 5.6, with summary statistics presented in Table 5.4. Sampling events prior to the addition of supplemental nutrients (e.g. S01 through S10) and additional operational changes (e.g. addition of filter aid polymer, temperature control, changes in chlorine quenching, and major fluctuations in targeted ozone doses) were evaluated as steady state operation. Sampling events after S10 were evaluated as a comparison to steady state operation and assessed the ability for bacterial microbiota to respond to changing conditions.

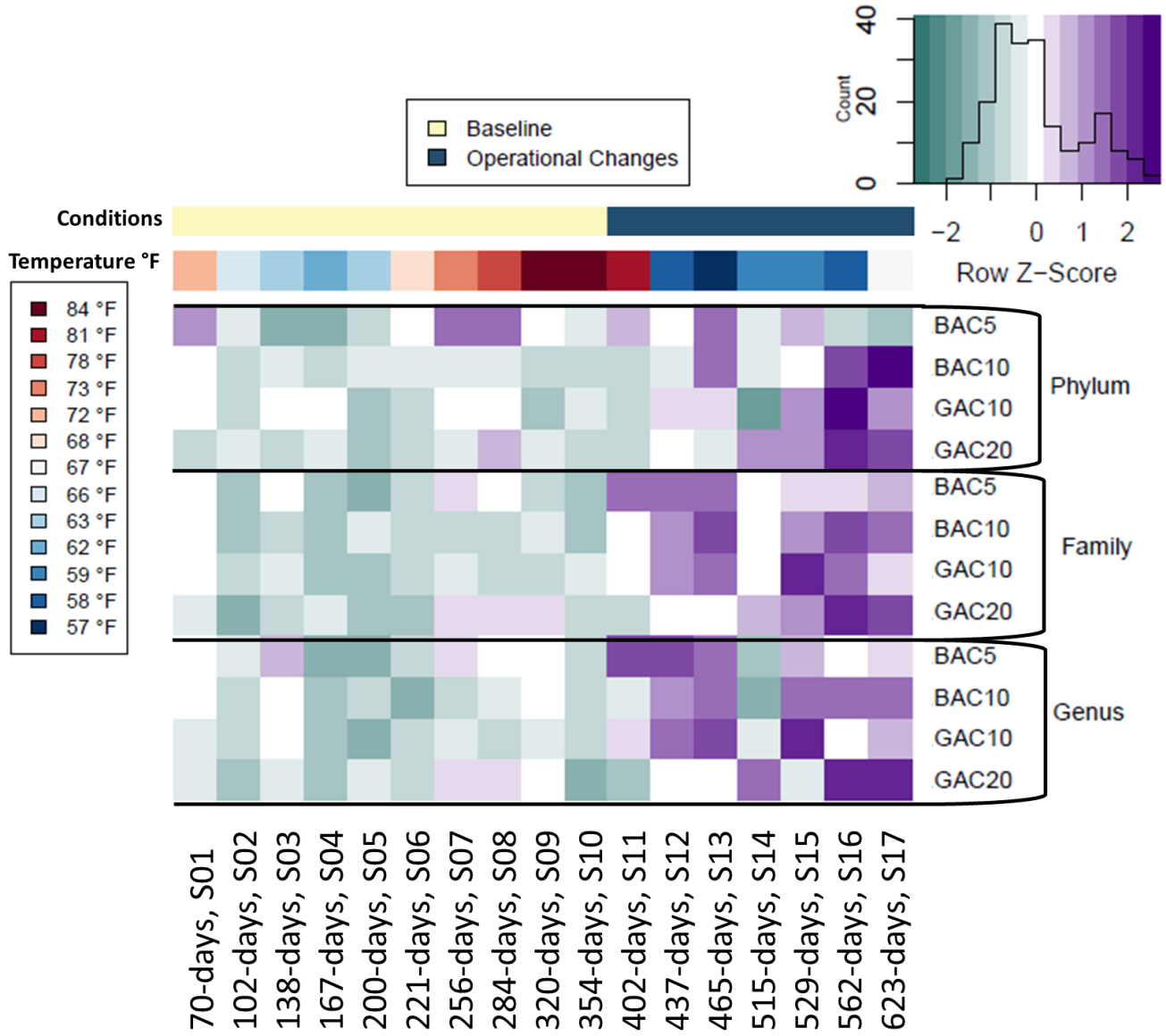


Figure 5.6: Succession in the biologically active filters with time, temperature, and treatment manipulation. Heatmap of z-scores associated with the average, cumulative temporal change in relative abundances of community taxa within the BAC filters and GAC contactors. Taxa are presented for each treatment process at the Phylum, Family, and Genus levels and are contextualized to changes in operation (baseline vs operational changes), and temperature.

BAC Filtration

During steady state operation, there was little change in the relative abundances of phyla represented in BAC5 (avg = 10%) and BAC10 (avg = 10%). Shifts in average relative abundances of taxa classified at the family and genus level were higher in magnitude and found to be consistent within individual EBCTs. Specifically, change in relative abundance averaged 38% at the family level and 41% at the genus level in BAC5, while BAC10 experienced, on average, changes of 29% and 35% at the family and genus levels, respectively. During this

period, changes were consistently elevated during the first sampling (change between S00 and S01), with BAC5 exhibiting increases in relative abundance during S03 (only family and genus classifications), S07, and S08 at all taxonomy levels analyzed. Comparatively, BAC10 exhibited slightly lower elevated changes only during S03, and primarily at the phylum and genus level.

During operation with conditional changes, the phylum level relative abundance turnover failed to increase within BAC5 (avg = 10%) with elevated turnover during BAC10 (avg = 33%). Interestingly, BAC5 and BAC10 both exhibited comparable increased turnover in their family and genus classified taxa.At the family level BAC5 averaged a 72% change in relative abundance compared to 70% at BAC10, with genus level changes averaging 68% and 66% for BAC5 and BAC10, respectively. Though similar in magnitude, BAC5 tended to exhibit higher changes in relative abundance immediately after the addition of supplemental nutrients and temperature control with slightly depressed changes in later sampling events with BAC10 experiencing the inverse.

Table 5.4: Summary Statistics for Temporal Change in Community Relative Abundance Across Multiple Taxonomic Classifications. Note: Values show the average percent change in classified reads between sampling events at each specified location and classification level.

Classification	Location	All Sampling Events		Before Operational Changes		After Operational Changes	
		Average Change in Realtive Abundance	Std	Average Change in Realtive Abundance	Std	Average Change in Realtive Abundance	Std
Phylum	BAC5	0.10	0.06	0.10	0.06	0.10	0.05
	BAC10	0.20	0.20	0.10	0.05	0.33	0.26
	GAC10	0.20	0.07	0.18	0.03	0.24	0.10
	GAC20	0.33	0.18	0.24	0.09	0.46	0.21
Family	BAC5	0.52	0.22	0.38	0.14	0.72	0.14
	BAC10	0.46	0.25	0.29	0.09	0.70	0.19
	GAC10	0.46	0.21	0.32	0.07	0.66	0.18
	GAC20	0.49	0.18	0.40	0.11	0.62	0.18
Genus	BAC5	0.52	0.24	0.41	0.17	0.68	0.25
	BAC10	0.48	0.25	0.35	0.12	0.66	0.28
	GAC10	0.48	0.19	0.37	0.08	0.65	0.18
	GAC20	0.47	0.20	0.39	0.12	0.59	0.24

GAC Contacting

During steady state operation, GAC10 experienced average changes in relative abundance at the phylum level ~18% while GAC20 experienced average changes of 24%, both of which were elevated to those experienced by both BACs during the same sampling period. At the family and genus level, both GAC contactors exhibited elevated and consistent changes to relative abundance ranging from 32% to 40%. Following changes to operation, increases in relative abundance were observed across all taxonomic levels, with greater changes in magnitude at the family and genus levels. Specifically, GAC10 was found to average a change of 24% at the phylum level with 46% average at GAC20. At the family and genus levels both GAC contactors consistently average changes ranging from 59% to 66%, which were comparable to changes identified during BAC treatment.

DISCUSSION

This study provides insight into the microbial community composition of a carbon-based potable water reuse treatment train and its succession in space and time as the water passes through a series of various physical, chemical, and microbiological treatments. While carbon-based treatment trains are being considered as an alternative to more costly membrane-based processes, it is important to develop a thorough understanding of the microbial populations involved. This is especially the case in potable reuse scenarios, where biological treatment has been slower to be embraced, in part due to concerns regarding uncertainty of the microbes involved, their capacity for removal of nutrients and trace organic contaminants, their resilience in the face of disturbance, and their overall biological safety. Generally, it was found that there is a distinct baseline microbiota associated with each stage of treatment that undergoes succession with time. Chemical treatments, such as O_3, resulted in the sharpest shifts, and also variance, in microbial community composition. Other operational changes, such as adjustment in O_3:TOC, temperature, filter aid polymer, chlorine quenching agent, and EBCT also resulted in measurable changes in the baseline microbial community composition of each process, but to a lesser degree. Of these, supplementation of nitrogen and phosphorus resulted in the strongest bifurcation and was particularly noted in the analysis microbial communities inhabiting the BAC and GAC units.

Implications of Physical Separation During Coagulation, Flocculation, and Sedimentation

The Floc-Sed process, at most, altered the relative abundance of a handful of taxa, but did not substantially shift the broader microbial community composition. Most of the variation was explained by individual sampling events with time, which is consistent with a purely physical separation in which upstream process variability was the main driver of observed variation in the effluent. Although temperature varied substantially throughout the study (it was not controlled at the front of the pilot), had a negligible effect.

Interestingly, most bacterial Families that decreased in the Floc-Sed unit were associated with known denitrifying taxa, which suggests loss of this niche following the upstream denitrifying filter. These include Methylophilaceae, which has been commonly identified in denitrifying biofilters in other studies and can utilize methanol as a sole carbon and energy source [31,32]. This is consistent with the addition of methanol to the upstream denitrifying filters, which depleted by the time the water reaches the Floc-Sed unit. Rhodocyclaceae, known to contain denitrifiers [33], Xanthomonadaceae, and Flavobacteriaceae also decreased through Floc-Sed treatment. Notably, Rhodocyclaceae, Xanthomonadaceae, and Flavobacteriaceae were found to be strongly associated with COD and TN removal within an aerobic granular sludge reactor treating anaerobic digestion wastewater [34]. Given that these organisms are reduced, but not eliminated, they could conceivably persist and aid with downstream carbon and nitrogen removal in the carbon-based train.

The Families that increased in relative abundance through the Floc-Sed were also logical, based on what is known about their associated functions. For example, Bdellovibrio and Bacteriovoracaceae are known to prey upon other bacteria [35–37], particularly Gram negative bacteria inhabiting biofilms. Further, bacteria known to depend on amoebae hosts for replication, such as Legionellaceae and Procabacteriaceae [38], also increased during Floc-Sed treatment. Given that predatory bacteria and amoebae can disrupt flocs, they could be of interest for future

study if settling problems arise. Effluent turbidity was identified as having the strongest relationships with these taxa.

Impact of Ozonation Volatility on Microbial Communities

Ozonation instigated the sharpest shift in microbial composition, which is logical given that acts as a disinfectant. Interestingly, ozonation also induced the greatest variance in microbial composition, which likely relates to its semi-indiscriminate oxidation of various taxa of microbes and their DNA. Like Floc-Sed, the composition was clearly shaped by time, as reflected by individual sampling events, which is consistent with non-specific chemical reactions paring back a broad range of incoming microbial populations. Factors that impact oxidative potency of ozone influenced the manner in which it shaped the resulting microbial community. Microbial community composition was clearly shaped by O_3: TOC ratio and was best explained when adjusted for nitrite concentration and temperature. This is logical given that nitrite is a sink for ozone, while its overall oxidative capacity is temperature dependent. Other ozone sinks, such as suspended solids and NOM, are also known to affect potency and would not be accounted for in solely considering O_3:TOC [39,40].

Ozonation was not fully indiscriminate in its action against various microbial populations. Gram negative bacteria appeared to be more susceptible, likely because their thinner cell walls afford them less protection against ozone. Accordingly, Legionellaceae and other families of amoebae-resisting bacteria were effectively reduced. Anaerobic bacteria also fared worse, which is logical given that post-ozonation dissolved oxygen levels can approach 35.9±9.5 mg/L. On the other hand, most of the organisms that were enriched by ozonation belonged to families associated with biofilm formation. This is also logical given that biofilms afford protection against disinfectants [41,42], including ozone [43]. Specifically, Hyphomonadaceae are known to produce holdfast, a polar polysaccharide that facilitates initial stages of biofilm formation [44]. Several species of Oxalobacteraceae are also known to form biofilms, especially under stress conditions [45–47]. Caulobacteraceae are known for their ability to form biofilms as a means of thriving in oligotrophic habitats [48] and are commonly found in drinking water and bottled water [49]. Nitrosomonadaceae, a Family of ammonia oxidizers, were also enriched post-ozonation following most sampling events, but was probably a result of the addition of preformed monochloramine during ozonation for bromate suppression. Interestingly, Comamonadaceae, a suspected bromate reducer [50–52], was also positively correlated with bromate concentrations and might also be aiding in suppressing bromate levels. Overall, the findings suggest that some populations of microbes not only survive, but through the shelter of biofilms, grow as the water passes through the ozone unit.

Microbial Dynamics of BAC Filtration for Contaminant Removal

BAC filters are the centerpiece of carbon-based potable water reuse treatment, serving as a hub for microbiological breakdown of trace organic compounds. Given that ozone acts to break-up large refractory organics into more biodegradable forms, it was of interest to determine how this was reflected in the resulting microbial populations. The microbial community analysis herein is consistent with heterotrophic generalists primarily colonizing the BAC filters, with the simple organic carbon compounds providing little incentive for the selection of microbes capable of degrading more recalcitrant compounds. Of the taxa that were consistently enriched throughout both BAC filters, only Hyphomicrobiaceae is well known for its involvement in

bioremediation [53]. Oxalobacteraceae were further enriched in BAC5 and are known for diverse metabolic functions, application for various biotechnological purposes, and resilience to environmental stress [46,54]. Opitutaceae were enriched in BAC10 and have been noted for their potential for bioremediation applications and are included in the phylum Verrucomicrobia, which are known to carry a diverse set of polysaccharide degradation enzymes [55–57]. Outside of Oxalobacteraceae and Pseudomonadaceae, no additional taxa were correlated with bulk TOC removal. The overall lack of selection of multiple taxa known for their capacity to degrade recalcitrant compounds by the BAC filters is consistent with on the observation that the majority of TOC removed by these units is readily biodegradable (Blair et al., in review).

Microbial community profiling pointed to substantial nitrogen cycling in the BAC units, especially in BAC10. Key Families that consistently increased in relative abundance in the BAC10 filter relative to the ozone effluent included Nitrosomonadaceae, Phyllobacteriaceae, and Planctomycetaceae, which contain members that oxidize ammonia [58], fix nitrogen [59], and anaerobically oxidize ammonia (anammox) [60], respectively.

The BAC filters also bore the brunt of key experimental changes to the system, and it was of further interest to assess the resilience of corresponding microbial populations. O_3:TOC ratio was under investigation during this study period and was found to correlate with the relative abundance of multiple Families. The introduction of nitrogen and phosphorus to the feed midway through the study was found to be especially impactful, leading to an across the board shift in bacterial community composition (Figure 5.4c). Families that increased in relative abundance in response to nutrient addition for both BAC filters included: Flavobacteriaceae, Sphingobacteriaceae, Sphingomonadaceae, and Hydrogenophilaceae. Nitropiraceae, a well-known nitrifier, was also enriched by the nutrient addition with Nitrosomonadaceae's selection increasing on BAC5 and remaining high on BAC10. Temperature, which is a key determinant of microbiological growth rate and enzyme activity, also had an effect. However, the effect was more notable prior to nutrient addition and before temperature control were commenced, during which time a wider range of temperatures were experienced. Changes in bacterial community structure were also associated with potential presence of chlorine in the non-quenched events, which is logical, given that it is a disinfectant. Thiosulfate and bisulfite addition could have shifted the community due to quenching of this disinfectant, but also because it can serve as an electron donor/acceptor and as a growth nutrient (Friedrich, 2001). The latter is consistent with Halothiobacillaceae, known for oxidation of thiosulfate [62,63], increasing after the switch to thiosulfate chlorine quenching. Filter aid polymer was found to have the least impact on BAC effluent communities, which is consistent with minimal impact observed for Floc-Sed as a physical process. In general, the effects of each manipulation were more apparent when analyzing a subset of data in which the other variables were unchanged. Notably, the BAC10 unit was consistently found to house a more stable taxonomic composition than the BAC5 unit, suggesting that a shorter EBCT may provide less resilience to changes, such as those imposed herein.

Colonization and Succession of Microbial Communities During GAC contactors

In contrast to the BAC filters, the GAC contactors primarily receive recalcitrant organics that were not readily degraded during BAC treatment. Consistent with the hypothesis that this would select for distinct metabolic capabilities, the microbial community compositions were found to differ between the GAC versus BAC units. Further, a greater number of taxa associated

with specialized metabolism or nutrient cycling were found to increase through the GAC units, including: Acetobacteraceae [64,65], Caldilineaceae [66], Cryomorphaceae [67,68], Opitutaceae, Planctomycetaceae, Rhodospirillaceae [69,70], Sphingobacteriaceae [71,72], and Nitrospiraceae [73]. Consistent with the upstream BAC process generating biomass and available carbon being scarce in the GAC contactors, indicators of predatory metabolism were also apparent (e.g., Coxiellaceae, Bdellobrioanaceae, and Bacteriovoracacae), as they were in the Floc-Sed unit. Overall, the findings emphasize that GAC contactors are not purely physical adsorption units, rather, they house a robust microbial community of their own that can potentially be harnessed to optimize removal of recalcitrant compounds. Future research focused on characterizing isolates from the GAC media or profiling their functional genes could provide further insight beyond that which was discernable from 16S rRNA gene analysis.

Also in contrast to the BAC filters, the GAC contactors experienced downstream, rather than direct, effects of experimental manipulations and thus a more stable influent. Correspondingly, colonization and succession dynamics were more stable during GAC contacting. Still, microbial community composition of the GAC units did shift in response to changes in operational variables (e.g. supplemental nutrients, variable ozone dosing, filter aid polymer, chlorine quenching, temperature) in a manner that mirrored the BAC filters. As was observed in the BAC filters, the taxonomic structure of the GAC contactors clearly shifted in response to nutrient addition, although to a lesser degree than was observed in the BAC filters. Notably, correspondence analysis revealed several Families in common between the BAC and GAC units during nutrient supplementation (i.e., Rhodocyclaceae, Erythrobacteraceae, Clostridiaceae, Verrucomicrobiaceae, Peptosteptococcaceae, Procabacteriaceae, Enterobacteriaceae, Thiotrichaceae, Halothiobacillaceae, and Hydrogenophilaceae), but only two before nutrient supplementation (Alcaligenaceae and Leptospiraceae). The observed shifts in GAC taxa also did not translate into any apparent correlations with changes in operation or measured parameters. In contrast to the BAC filters, Nitrosomonadaceae were unimpacted by the ammonia addition, indicating that most ammonia oxidation occurring prior to GAC contacting, which was consistent with measured ammonia concentrations.

UV Disinfection

Among all processes employed across the treatment train, UV disinfection generally resulted in the least explainable change in microbial community composition by factors tested. This is consistent with UV primarily functioning as a physical process that inactivates bacteria, but does not markedly shift the proportion of DNA originating from different microbes captured in corresponding sample analysis. Overall, UV disinfection was found to reduce the relative abundance of most abundant bacterial taxa, with only a few examples of consistent positive increases. However, relative abundances are a zero-sum measure, and reductions in one taxa's abundance must be accounted for in an increase of another. In this way, taxa can experience relative abundance increases even when the individual relative abundance decreased via a disproportionate reduction when compared to other taxa. Mycobacteriaceae's consistent increase in relative abundance is especially noteworthy because some members are human pathogens [74]. The finding is consistent with extreme resistance of Mycobacteriaceae to UV [75,76]. Few comparable studies of UV effects on microbial community composition were found in the literature, as most prior studies have focused on either wastewater or drinking water. In one wastewater study, Comamonadaceae, Moraxellaceae, and Rhodocyclaceae were similarly found

to be enriched by UV [77]. Moraxellaceae has been noted to be UV tolerant [78]. Comparison to a drinking water-based study jointly identified phylum-level trends, with Bacteroidetes being susceptible to UV treatment and Actinobacteria (Gram positive organisms) being more resistant [75]. Sphingomonadaceae was also consistently enriched by UV and has been identified in several studies for its ability to produce EPS and form biofilms [79,80] (which could potentially foul UV units) as well as degrade xenobiotic and recalcitrant (poly)aromatic compounds [81,82]. Sphingomonadaceae has also been identified as an early colonizer of reverse osmosis and nanofiltration membranes that leads to biofouling [83].

The lack of meaningful correlations to operational conditions and environmental parameters, and the limited consistently positive selection of taxa tend to indicate that the DNA markers that persist through UV treatment are likely the result of several compounded factors. These can include changes in water chemistry, microbial aggregate and interactions, the nature of the cell envelope, inherent or acquired UV resistance, more efficient regrowth following treatment, and/or activation of UV repair mechanisms [75].

Stability and Resilience of Carbon-Based Advanced Water Treatment

Evaluation of the BAC and GAC units during periods of stable operation as well as periods where conditions were changing provided insight into the stability of the microbial communities and their resilience in the face of disturbance. Notably, even during stable operation, consistent succession (ranging from 10-41% cumulative change in relative abundances) was observed in the BAF units across all taxonomic levels analyzed, with greater turnover at the family and genus levels than the phylum level. Major shifts in operation correspondingly led to increased average succession rates, ranging from 10-72%. Thus, the analysis approach demonstrated herein helps to establish the baseline turnover of microbial communities in carbon-based advanced water treatment, which then can serve as a reference point for changes that occur over and above this level in response to specific intentional, or unintentional, operational changes.

Interestingly, temporal proximity to the most recent backwashing event did not correspond well to changes in the overall microbial community structure. This suggests relative resilience of the carbon-based train to backwashing. On the other hand, S01 was noted to have an elevated succession rate, which is likely attributed to the initial acclimation period associated with biofilter operation, where the virgin GAC was initially colonized and the previously exhausted BAC filters (on a difference source water) experienced a significant influent change. Following initial acclimation, S07 and S08 aligned with transfer of 5 inches of media was from BAC10 into BAC5. Correspondingly, S07 and S08 exhibited much higher succession percentages for BAC5 than BAC 10, and compared to previous sampling dates. Further, there was an appreciable increase in TOC removal following the inoculation. These findings suggest that it would be worthwhile for future studies to assess the potential for inoculation to enhance performance.

EBCT was also of interest in this study, as shorter EBCTs could increase the overall treatment rate. A longer EBCT may be necessary, if maximizing removal of recalcitrant organics is the concern. In terms of taxonomic composition, phylum-level microbial succession rates were higher at higher EBCTs in both BAC and GAC systems (outside of baseline conditions), suggesting large-scale differences in microbiota might be driven by EBCT. On the

other hand, EBCT under 5mins prevented greater phylum level turnover, regardless of operational conditions. In general, elevated EBCT appears to be more beneficial to the GAC unit, permitting increased opportunity for biodegradation and sorption of recalcitrant organics, following removal of simple organics by the BAC unit. On the other hand, there was more turnover in the BAC5 unit, suggesting there might be a trade-off with less resilience of BAC filtration and shorter EBCTs.

ACKNOWLEDGEMENTS

The authors would like to thank the operators and staff at Hampton Road Sanitation District's York River Treatment Plant and the SWIFT Research Center. Funding for this effort was provided in part by Water Research Foundation Unsolicited Award U1R16 (PI Pruden) and US Bureau of Reclamation grant R21AC10162 (PI Pruden) 09/01/2020-08/31/2023 leveraged with additional financial and in-kind support from the Hampton Roads Sanitation District. Special thanks to Rachel Craine and Pruden Lab members for assistance in the laboratory.

REFERENCES

1. Kalkan, Ç., Yapsakli, K., Mertoglu, B., Tufan, D. & Saatci, A. Evaluation of Biological Activated Carbon (BAC) process in wastewater treatment secondary effluent for reclamation purposes. *Desalination* **265**, 266–273 (2011).

2. Plumlee, M. H., Stanford, B. D., Debroux, J. F., Hopkins, D. C. & Snyder, S. A. Costs of Advanced Treatment in Water Reclamation. *http://dx.doi.org/10.1080/01919512.2014.921565* **36**, 485–495 (2014).

3. Garner, E. *et al.* Next generation sequencing approaches to evaluate water and wastewater quality. *Water Res.* **194**, 116907 (2021).

4. Shokralla, S., Spall, J. L., Gibson, J. F. & Hajibabaei, M. Next-generation sequencing technologies for environmental DNA research. *Mol. Ecol.* **21**, 1794–1805 (2012).

5. Boon, N., Pycke, B. F. G., Marzorati, M. & Hammes, F. Nutrient gradients in a granular activated carbon biofilter drives bacterial community organization and dynamics. *Water Res.* **45**, 6355–6361 (2011).

6. Feng, S. *et al.* Characterization of microbial communities in a granular activated carbon-sand dual media filter for drinking water treatment. *Int. J. Environ. Sci. Technol.* **10**, 917–922 (2013).

7. Lautenschlager, K. *et al.* Abundance and composition of indigenous bacterial communities in a multi-step biofiltration-based drinking water treatment plant. *Water Res.* **62**, 40–52 (2014).

8. Liao, X. *et al.* Changes of biomass and bacterial communities in biological activated carbon filters for drinking water treatment. *Process Biochem.* **48**, 312–316 (2013).

9. Stoquart, C., Barbeau, B., Servais, P. & Vázquez-Rodríguez, G. A. Quantifying bacterial biomass fixed onto biological activated carbon (PAC and GAC) used in drinking water treatment. *J. Water Supply Res. Technol.* **63**, 1–11 (2014).

10. Balcioğlu, I. A. & Ötker, M. Treatment of pharmaceutical wastewater containing antibiotics by O3 and O3/H2O2 processes. *Chemosphere* **50**, 85–95 (2003).

11. Lohwacharin, J., Phetrak, A., Takizawa, S., Kanisawa, Y. & Okabe, S. Bacterial growth during the start-up period of pilot-scale biological activated carbon filters: Effects of residual ozone and chlorine and backwash intervals. *Process Biochem.* **50**, 1640–1647 (2015).

12. Gibert, O. *et al.* Characterising biofilm development on granular activated carbon used for drinking water production. *Water Res.* **47**, 1101–1110 (2013).

13. Magic-Knezev, A. & van der Kooij, D. Optimisation and significance of ATP analysis for measuring active biomass in granular activated carbon filters used in water treatment. *Water Res.* **38**, 3971–3979 (2004).

14. Reungoat, J. *et al.* Ozonation and biological activated carbon filtration of wastewater treatment plant effluents. *Water Res.* **46**, 863–872 (2012).

15. Moll, D. M., Summers, R. S., Fonseca, A. C. & Matheis, W. Impact of Temperature on Drinking Water Biofilter Performance and Microbial Community Structure. *Environ. Sci. Technol.* **33**, 2377–2382 (1999).

16. Fonseca, A. C., Scott Summers, R. & Hernandez, M. T. Comparative measurements of microbial activity in drinking water biofilters. *Water Res.* **35**, 3817–3824 (2001).

17. Boon, E. *et al.* Interactions in the microbiome: communities of organisms and communities of genes. *FEMS Microbiol. Rev.* **38**, 90–118 (2014).

18. Gerrity, D. *et al.* Microbial community characterization of ozone-biofiltration systems in drinking water and potable reuse applications. *Water Res.* (2018) doi:10.1016/j.watres.2018.02.023.

19. Stamps, B. W. *et al.* Characterization of the Microbiome at the World's Largest Potable Water Reuse Facility. *Front. Microbiol.* **0**, 2435 (2018).

20. Pruden, A., Bott, C., Blair, M. F., Miller, J. H. & Vaidya, R. Characterization of Organic Carbon and Microbial Communities for the Optimization of Biologically-Active Carbon (BAC) Filtration for Potable Reuse. *Water Res. Found. Proj. 4872/U1R16 Final Report.* 305-undefined (2020).

21. Sun, Y. *et al.* A pilot-scale investigation of disinfection by-product precursors and trace organic removal mechanisms in ozone-biologically activated carbon treatment for potable reuse. *Chemosphere* **210**, 539–549 (2018).

22. Vaidya, R. *et al.* Pilot Plant Performance Comparing Carbon-Based and Membrane-Based Potable Reuse Schemes. *Environ. Eng. Sci.* **36**, 1369–1378 (2019).

23. Vaidya, R., Wilson, C. A., Salazar-Benites, G., Pruden, A. & Bott, C. Implementing Ozone-BAC-GAC in potable reuse for removal of emerging contaminants. *AWWA Water Sci.* **2**, e1203 (2020).

24. Vaidya, R., Wilson, C. A., Salazar-Benites, G., Pruden, A. & Bott, C. Factors affecting removal of NDMA in an ozone-biofiltration process for water reuse. *Chemosphere* **264**, 128333 (2021).

25. Bolyen, E. *et al.* Author Correction: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2 (Nature Biotechnology, (2019), 37, 8, (852-857), 10.1038/s41587-019-0209-9). *Nature Biotechnology* vol. 37 1091 (2019).

26. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).

27. Team, Rs. Citing RStudio – RStudio Support. *RStudio: Integrated Development for R. RStudio* https://support.rstudio.com/hc/en-us/articles/206212048-Citing-RStudio (2020).

28. Wickham, H. ggpolt2 Elegant Graphics for Data Analysis. *Use R! Ser.* 211 (2016).

29. Oksanen, J. *et al.* Package 'vegan' Title Community Ecology Package Version 2.5-7. (2020).

30. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).

31. Doronina, N., Kaparullina, E. & Trotsenko, Y. The family Methylophilaceae. *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* **9783642301971**, 869–880 (2014).

32. Rissanen, A. J., Ojala, A., Fred, T., Toivonen, J. & Tiirola, M. Methylophilaceae and Hyphomicrobium as target taxonomic groups in monitoring the function of methanol-fed denitrification biofilters in municipal wastewater treatment plants. *J. Ind. Microbiol. Biotechnol.* **44**, 35–47 (2017).

33. Wang, Z., Li, W., Li, H., Zheng, W. & Guo, F. Phylogenomics of Rhodocyclales and its distribution in wastewater treatment systems. *Sci. Reports 2020 101* **10**, 1–12 (2020).

34. Xiong, W. *et al.* High-strength anaerobic digestion wastewater treatment by aerobic granular sludge in a step-by-step strategy. *J. Environ. Manage.* **262**, 110245 (2020).

35. Jurkevitch, E. The ecology of Bdellovibrio and like organisms in wastewater treatment plants. in *The Ecology of Predation at the Microscale* 37–64 (Springer, Cham, 2020). doi:10.1007/978-3-030-45599-6_2.

36. Davidov, Y. & Jurkevitch, E. Diversity and evolution of Bdellovibrio-and-like organisms (BALOs), reclassification of Bacteriovorax starrii as Peredibacter starrii gen. nov., comb. nov., and description of the Bacteriovorax-Peredibacter clade as Bacteriovoracaceae fam. nov. *Int. J. Syst. Evol. Microbiol.* **54**, 1439–1452 (2004).

37. Kuppardt-Kirmse, A. & Chatzinotas, A. Intraguild predation: Predatory networks at the microbial scale. *Ecol. Predation Microscale* 65–87 (2020) doi:10.1007/978-3-030-45599-6_3/COVER.

38. Heinz, E. *et al.* An Acanthamoeba sp. containing two phylogenetically different bacterial endosymbionts. *Environ. Microbiol.* **9**, 1604 (2007).

39. Alvares, A. B. C., Diaper, C. & Parsons, S. A. Partial Oxidation by Ozone to Remove Recalcitrance from Wastewaters - a Review. *http://dx.doi.org/10.1080/09593332208618273* **22**, 409–427 (2010).

40. Gerrity, D. & Snyder, S. Review of Ozone for Water Reuse Applications: Toxicity, Regulations, and Trace Organic Contaminant Oxidation. *http://dx.doi.org/10.1080/01919512.2011.578038* **33**, 253–266 (2011).

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

42. Sanchez-Vizuete, P., Orgaz, B., Aymerich, S., Le Coq, D. & Briandet, R. Pathogens protection against the action of disinfectants in multispecies biofilms. *Front. Microbiol.* **6**, 705 (2015).

43. Kotlarz, N. *et al.* Biofilms in Full-Scale Drinking Water Ozone Contactors Contribute Viable Bacteria to Ozonated Water. *Environ. Sci. Technol.* **52**, 2618–2628 (2018).

44. Abraham, W. R. & Rohde, M. The family Hyphomonadaceae. *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* **9783642301971**, 283–299 (2014).

45. Gong, M., Yang, G., Zhuang, L. & Zeng, E. Y. Microbial biofilm formation and community structure on low-density polyethylene microparticles in lake water microcosms. *Environ. Pollut.* **252**, 94–102 (2019).

46. Baldani, J. I. *et al.* The family Oxalobacteraceae. *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* **9783642301971**, 919–974 (2014).

47. Nguyen, N. H. A. *et al.* Early stage biofilm formation on bio-based microplastics in a freshwater reservoir. *Sci. Total Environ.* **858**, 159569 (2023).

48. Abraham, W. R., Rohde, M. & Bennasar, A. The family Caulobacteraceae. *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* **9783642301971**, 179–205 (2014).

49. Gonzalez, C., Gutierrez, C. & Grande, T. Bacterial flora in bottled uncarbonated mineral drinking water. *https://doi.org/10.1139/m87-196* **33**, 1120–1125 (2011).

50. Liu, J., Yu, J., Li, D., Zhang, Y. & Yang, M. Reduction of bromate in a biological activated carbon filter under high bulk dissolved oxygen conditions and characterization of bromate-reducing isolates. *Biochem. Eng. J.* **65**, 44–50 (2012).

51. Lai, C. Y. *et al.* Bromate and Nitrate Bioreduction Coupled with Poly-β-hydroxybutyrate Production in a Methane-Based Membrane Biofilm Reactor. *Environ. Sci. Technol.* **52**, 7024–7031 (2018).

52. Wang, D. *et al.* Phylogenetic characterization of bromate-reducing microbial community enriched anaerobically from activated sludge. *Ecotoxicol. Environ. Saf.* **184**, 109630 (2019).

53. Oren, A. & Xu, X. W. The family hyphomicrobiaceae. *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* **9783642301971**, 247–281 (2014).

54. Zhang, Y. & He, Q. Characterization of bacterial diversity in drinking water by pyrosequencing. *Water Supply* **13**, 358–367 (2013).

55. Rodrigues, J. L. M. & Isanapong, J. The family opitutaceae. *Prokaryotes Other Major Lineages Bact. Archaea* **9783642389542**, 751–756 (2014).

56. Martinez-Garcia, M. *et al.* Capturing Single Cell Genomes of Active Polysaccharide Degraders: An Unexpected Contribution of Verrucomicrobia. *PLoS One* **7**, e35314 (2012).

57. Lin, Z. *et al.* Enhancing pentachlorophenol degradation by vermicomposting associated bioremediation. *Ecol. Eng.* **87**, 288–294 (2016).

58. Prosser, J. I., Head, I. M. & Stein, L. Y. The family Nitrosomonadaceae. *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* **9783642301971**, 901–918 (2014).

59. Willems, A. The family Phyllobacteriaceae. in *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* vol. 9783642301 355–418 (2014).

60. Kuenen, J. G. Anammox bacteria: From discovery to application. *Nat. Rev. Microbiol.* **6**, 320–326 (2008).

61. Friedrich, C. G., Rother, D., Bardischewsky, F., Quentmeier, A. & Fischer, J. Oxidation of Reduced Inorganic Sulfur Compounds by Bacteria: Emergence of a Common Mechanism? *Applied and Environmental Microbiology* vol. 67 2873–2882 (2001).

62. Boden, R. Reclassification of Halothiobacillus hydrothermalis and Halothiobacillus halophilus to Guyparkeria gen. Nov. in the Thioalkalibacteraceae fam. nov., with emended descriptions of the genus Halothiobacillus and family Halothiobacillaceae. *Int. J. Syst. Evol. Microbiol.* **67**, 3919–3928 (2017).

63. Vannini, C. *et al.* Sulphide oxidation to elemental sulphur in a membrane bioreactor: Performance and characterization of the selected microbial sulphur-oxidizing community. *Syst. Appl. Microbiol.* **31**, 461–473 (2008).

64. Komagata, K., Iino, T. & Yamada, Y. The family Acetobacteraceae. *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* **9783642301971**, 3–78 (2014).

65. Covarrubias-García, I., de Jonge, N., Arriaga, S. & Nielsen, J. L. Effects of ozone treatment on performance and microbial community composition in biofiltration systems treating ethyl acetate vapours. *Chemosphere* **233**, 67–75 (2019).

66. Zhang, B., Xu, X. & Zhu, L. Structure and function of the microbial consortia of activated

sludge in typical municipal wastewater treatment plants in winter. *Sci. Reports 2017 71* **7**, 1–11 (2017).

67. Bowman, J. P. Out From the Shadows – Resolution of the Taxonomy of the Family Cryomorphaceae. *Front. Microbiol.* **11**, (2020).

68. Bowman, J. P. The family cryomorphaceae. *Prokaryotes Other Major Lineages Bact. Archaea* **9783642389542**, 539–550 (2014).

69. Baldani, J. I. *et al.* The family Rhodospirillaceae. *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* **9783642301971**, 533–618 (2014).

70. Madigan, M., Cox, S. S. & Stegeman, R. A. Nitrogen fixation and nitrogenase activities in members of the family Rhodospirillaceae. *J. Bacteriol.* **157**, 73–78 (1984).

71. Lambiase, A. The family sphingobacteriaceae. *Prokaryotes Other Major Lineages Bact. Archaea* **9783642389542**, 907–914 (2014).

72. Mao, J., Luo, Y., Teng, Y. & Li, Z. Bioremediation of polycyclic aromatic hydrocarbon-contaminated soil by a bacterial consortium and associated microbial community changes. *Int. Biodeterior. Biodegradation* **70**, 141–147 (2012).

73. Daims, H. The family nitrospiraceae. *Prokaryotes Other Major Lineages Bact. Archaea* **9783642389542**, 733–749 (2014).

74. González-Martínez, A. *et al.* Mycobacterium infections. *Encycl. Infect. Immun.* 703–718 (2022) doi:10.1016/B978-0-12-818731-9.00216-0.

75. Pullerits, K. *et al.* Impact of UV irradiation at full scale on bacterial communities in drinking water. *npj Clean Water 2020 31* **3**, 1–10 (2020).

76. Shin, G. A., Lee, J. K., Freeman, R. & Cangelosi, G. A. Inactivation of Mycobacterium avium complex by UV irradiation. *Appl. Environ. Microbiol.* **74**, 7067–7069 (2008).

77. Kauser, I., Ciesielski, M. & Poretsky, R. S. Ultraviolet disinfection impacts the microbial community composition and function of treated wastewater effluent and the receiving urban river. *PeerJ* **2019**, (2019).

78. Hare, J. M., Bradley, J. A., Lin, C. L. & Elam, T. J. Diverse responses to UV light exposure in Acinetobacter include the capacity for DNA damage-induced mutagenesis in the opportunistic pathogens Acinetobacter baumannii and Acinetobacter ursingii. *Microbiology* **158**, 601–611 (2012).

79. Glaeser, S. P. & Kämpfer, P. The family Sphingomonadaceae. *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* **9783642301971**, 641–707 (2014).

80. Koskinen, R. *et al.* Characterization of Sphingomonas isolates from Finnish and Swedish drinking water distribution systems. *J. Appl. Microbiol.* **89**, 687–696 (2000).

81. Oh, S. & Choi, D. Microbial Community Enhances Biodegradation of Bisphenol A Through Selection of Sphingomonadaceae. *Microb. Ecol.* **77**, 631–639 (2019).

82. Stolz, A. Molecular characteristics of xenobiotic-degrading sphingomonads. *Appl. Microbiol. Biotechnol.* **81**, 793–811 (2009).

83. de Vries, H. J. *et al.* Isolation and characterization of Sphingomonadaceae from fouled membranes. *npj Biofilms Microbiomes 2019 51* **5**, 1–9 (2019).

CHAPTER 6: WHAT IS THE DIFFERENCE BETWEEN CONVENTIONAL DRINKING WATER, POTABLE REUSE WATER, AND NON-POTABLE REUSE WATER? A CORE MICROBIOME PERSPECTIVE

Matthew F. Blair, Emily Garner, Pan Ji, and Amy Pruden

ABSTRACT

Motivations toward expanded implementation of a fit-for-(re)use framework and expansion of next generation sequencing technologies' applications have inspired an increasing number of studies aimed at characterizing the microbial communities inhabiting various water environments. Microbiome analysis could provide the ability to move beyond simplistic fecal indicator and pathogen profiling, towards defining a "healthy" drinking water microbiota. Here we utilized 16S rRNA gene amplicon sequencing to identify signature features in the composition of bacterial microbiota across a wide spectrum of water types (potable, potable reuse, and non-potable reuse). A clear distinction was found in the composition of bacterial microbiota as a function of intended water use across a very broad range of U.S. water systems at both the point of compliance (ANOSIM, r-stat=0.75) and point of use (ANOSIM, r-stat=0.34). Core and discriminatory analysis were also able to identify distinct differences between potable and non-potable water microbiomes at both the Phylum (OD1, OP3, Chlorobi, Verrucomicrobia and GN02) and Genus (*Arcobacter*, *Fluviicola, and Azospira*) levels. The approach and findings of this study open the door to the possibility of microbial community signature profiling as a water quality monitoring approach for assessing efficacy of treatments and suitability of water for intended use.

INTRODUCTION

Drinking water and wastewater systems provide rich environments that support taxonomically and functionally diverse microbial populations [1–4]. In the wastewater sector, biological treatment is essential for the removal of biochemical oxygen demand (BOD) and for nitrogen [5–7] and phosphorus [8–12] transformation and removal. Microbes in wastewater also contribute to degradation of contaminants of emerging concern (CECs) [13–17]. In the drinking water sector, microorganisms can also contribute to carbon and nutrient reduction [18,19], as well as microbial-induced corrosion [20,21] and waterborne disease [22,23]. Pathogenic microorganisms are of high concern in any water system. Regardless of where on the "one water" continuum[24] a given treatment or water application may fall, microbial community composition is a key driver of the realization of water quality goals [25] . As the water and wastewater industries move toward a fit for (re)use framework, systematic profiling of microbial communities could provide a comprehensive and high-resolution approach to water quality monitoring.

Advances in next generation sequencing (NGS) technologies have inspired an increasing number of studies aimed at characterizing the microbial communities, herein referred to as microbiomes, inhabiting various water environments. NGS has now widely been applied to profile microbes involved in drinking water and wastewater treatment unit processes [4,26–29], including assessing the impacts of operational conditions, such as: treatment technology employed [30], distribution system characteristics (e.g. pipe materials, water age, and water chemistry) [31], disinfectant types and concentrations [32], and seasonal and source water variation

[33,34]. Others have examined the microbial communities that colonize non-potable reuse water systems [35,36], various strategies for potable reuse (such as blending) [37], membrane filtration [38], and carbon-based biofiltration [30], as well as more conventional drinking water treatment plants and distribution systems [39]. However, such studies have largely been conducted in isolation, without direct comparison to identify shared versus distinct features represented across respective microbiomes. Such knowledge could help to define microbiome "signatures" as an alternative, high-resolution means of characterizing different water types. Two complimentary avenues of potentially achieving this are core and discriminatory microbiome analyses.

A core microbiome consists of organisms that are found in common across a given set of microbiomes and are hypothesized to play a key role within the ecosystem of interest [40]. The discriminatory microbiome, on the other hand, is made up of organisms that distinguish multiple ecosystems of interest [40]. Previously, core microbiome analysis has been applied towards comprehensively characterizing and comparing human microbiomes, specifically through the human microbiome project [41,42] and focused on important human niches, such as: oral microbiome [43], fecal microbiome [44], and gut microbiome [45,46]. To a lesser extent, core microbiome analysis has been applied to understand the microbial dynamics and composition occurring within the water treatment microbiome [39,47] and wastewater microbiome [48,49]. As NGS is beginning to become more widely applied in the water sector [50], there is a need to establish methodology to support systematic comparisons across waters of different intended end uses. For example, microbiome analysis could provide high resolution analysis across the spectrum of non-potable and potable reuse waters. Further, microbiome analysis could take a key step in moving beyond simplistic fecal indicator and pathogen profiling, towards defining a "healthy" drinking water microbiota.

Here we utilized 16S rRNA gene amplicon sequencing to identify signature features in the composition of bacterial microbiota across a spectrum of water types [potable (bottled water, conventional drinking water), potable reuse (direct potable reuse, indirect potable reuse), and non-potable reuse]. Specifically, the objectives of this study were to: 1) identify the occurrence of differential microbial 'signatures' related to intended water use, 2) utilize core and discriminatory analysis to identify taxa and/or OTUs that are specific to the various water types and assess the corresponding variance observed for each water type, and 3) assess shifts in key taxa as a function of various treatments, stages of treatment, disinfection techniques, stage of conveyance (point of compliance versus point of use), climate, and other variables. The approach and findings of this study open the door to the possibility of microbial community signature profiling as a water quality monitoring approach for assessing efficacy of treatments and suitability of water for intended use.

METHODS

Site description, sample collection, and preservation

A large archive of in-house samples and data sets collected across several prior studies representing a wide range of water types, treatment technologies, climates and regional locations [37,39,51] was examined (Table 6.1). Additional samples were collected for this study from various brands of off-the-shelf bottled water and over a period of 23 months from a carbon-based indirect potable reuse (IPR) train taking secondary effluent from a nearby treatment plant and treating it with coagulation-flocculation-sedimentation, ozone advanced oxidation (O_3),

biologically-active carbon filtration (BAC), granular activated carbon filtration (GAC), and UV disinfection. Pilot influent samples, representative of a non-potable reuse water, and pilot effluent samples, representative of an IPR water, were included in this comparison study. Samples were classified as being representative of either the point of compliance (POC), i.e., directly after final treatment and prior to distribution, or the point of use (POU), i.e., after being conveyed through a distribution system and/or building plumbing.

Table 6.1: Key Attributes of Water Systems and Samples Included in this Study

Comparison	Attributes	Potable			Potable Reuse			Non-Potable Reuse			Totals
		POU (n)	POC (n)	All (n)	POU (n)	POC (n)	All (n)	POU (n)	POC (n)	All (n)	
Water Use		132	63	244*	26	27	53	95	82	177	425
Climate	**BSh – hot semi-arid**	7	2	9				16	3	19	28
	BWk – cold desert	2	1	3	8	4	12				15
	Cfa – humid subtropical	56	22	78		15	15	39	68	107	200
	Csa – hot-summer Mediterranean	4	2	6	6	2	8				14
	Csb – warm-summer Mediterranean	33	6	39	12	6	18	40	11	51	108
	Dfa – hot-summer humid continental	10	10	20							20
	Dfb – warm-summer humid continental	20	20	40							40
Region	**Midwest**	30	30	60							60
	Northeast	10	10	20							20
	Southeast	46	12	58		15	15	39	68	107	180
	Southwest	23	4	27				36	11	47	74
	West	23	7	30	26	12	38	20	3	23	91
Disinfection Residual Type	**NH_2Cl**	54	16	70	18	8	26	95	26	121	217
	Cl_2	78	47	125	8	4	12				137
	None					15	15		56	56	71
DNA Extraction Kit	**FastDNA Spin Kit**	132	63	244	26	12	38	95	26	121	354
	FastDNA Spin Kit for Soil					15	15		56	56	71
Source Water	**Groundwater**	4	2	6	8	3	11				17
	Surface Water	56	53	109	18	9	27				136
	Surface and Groundwater	72	8	80							80
	Treated Wastewater					15	15	95	82	177	192
Disinfection Processes	**NH_2Cl**	38	3	41				16	3	19	60
	NH_2Cl O_3 Industry				2	1	3				3
	NH_2Cl O_3 UV					15	15				15
	NH2Cl UV	10	10	20	2		2		4	4	26
	Cl_2	54	36	90	4	2	6				96
	Cl_2 NH_2Cl							59	19	78	78
	Cl_2 NH_2Cl | UV NH_2Cl**							20		20	20

	Cl_2 NH_2Cl O_3	2	1	3							3
	Cl_2 NH_2Cl O_3 UV				2	1	3				3
	Cl_2 ClO_2	16	2	18							18
	Cl_2 O_3	2	1	3	6	3	9				12
	Cl_2 O_3 UV				2		2				2
	Cl2 UV	10	10	20	4	2	6				26
	Cl_2 UV Pasteurization				2	1	3				3
	UV				2	2	4				4
Treatment Processes	**Aeration Cl_2**	36	2	38							38
	Bardenpho Cl_2							16	10	26	26
	Bardenpho Cl_2 | Biofiltration IFAS UV **							20		20	20
	Biofiltration IFAS UV								4	4	4
	CAS Denit								56	56	56
	CAS Denit Cl_2							23	6	29	29
	CAS Denit Coag Floc Sed O_3 BAC GAC UV					15	15				15
	CAS DualMediaFiltration Cl_2							20	3	23	23
	CAS DualMediaFiltration|MBR NH_2Cl **							16	3	19	19
	CAS MF RO UV				2	2	4				4
	CAS MF RO UV O_3 Coag Floc Sed Filtration Cl_2				2		2				2
	CAS Nit PDenit BioP Coag Floc Sed Filtration Cl_2				4	2	6				6
	CAS O_3 BAF Coag Floc Sed Filtration Cl_2				4	2	6				6
	CAS O_3 BAF O_3 Coag Floc Sed Filtration Cl_2				2	1	3				3
	Cl_2	20	4	24							24
	Coag Floc Sed Filtration Cl_2	18	3	21							21
	Coag Sed Filtration Cl_2	20	20	40							40
	Coag Sed Filtration GAC UV Cl_2	10	10	20							20
	Coag Sed Filtration UF Cl_2	10	10	20							20
	Coag Sed Filtration UV NH_2Cl	10	10	20							20
	FE&MN Cl_2	2	1	3							3
	FE&MN Cl_2 CAS MF RO UV Cl_2				4	2	6				6

	FE&MN Cl_2 CAS UF RO UV Past Cl_2				2	1	3				3
	NH_2Cl	2	1	3							3
	NH_2Cl CAS MF RO UV				2		2				2
	O_3 Coag Filtration NH_2Cl O_3 Coag Floc Sed Filtration Cl_2	2	1	3							3
	O_3 Coag Filtration NH_2Cl CAS MF RO UV O_3 Coag Floc Sed Filtration Cl_2				2	1	3				3
	O_3 Coag Filtration NH_2Cl Industry				2	1	3				3
	O_3 Coag Floc Sed Filtration Cl_2	2	1	3							3
Reuse Blending Ratio	**0.05**				2	1	3				3
	0.1				16	7	23				23
	0.5				8	4	12				12
	None	132	63	244*	0	15	15	95	82	177	387

* Includes 49 bottled water samples ** Indicates mixing of two treatment trains
Köppen classification: Bsh – midlatitude steppe and desert; Bwk – cold desert; Cfa – humid subtropical; Csa hot-summer Mediterranean; Csb – warm-summer Mediterranean; Dfa – hot-summer humid continental; Dfb – warm-summer humid continental

All samples were collected with the same QA/QC standards, i.e., in sterile 1-L polypropylene bottles, transported to the laboratory at Virginia Tech on ice and processed within 24 h of collection. Samples were concentrated onto 0.22-µm mixed cellulose esters membrane filter (Millipore, Billerica, MA) before being fragmented with sterilized tweezers and stored at -20 °C. Filters were subject to extraction using either a FastDNA SPIN Kit or FastDNA SPIN Kit for Soil (MP Biomedicals, Solon, OH) according to the manufacturer's instructions. DNA extracts were stored at -20 °C (short term), or -80 °C (long term) prior to downstream molecular analysis.

16S rRNA Gene Amplicon Sequencing

Given that this was a study of archival data, there were some differences in the sequencing methodologies applied across the samples. Approximately three quarters of the samples [37,39,51] (i.e., the potable water, direct potable reuse water, and non-potable reuse water) were subjected to PCR amplification using the universal bacterial/archaeal primer set 515f/806r targeting the V4 region of the 16S rRNA gene [52]. The remaining quarter of the samples, (i.e., the bottled water, indirect potable reuse water, and additional non-potable reuse waters) were subjected to bar-coded PCR amplification using the modified universal bacterial/archaeal primer set 515f/926r targeting the V4-V5 region of the 16S rRNA gene [53]. Positive PCR products with corresponding negative control PCR for the run and passing QA/QC were Qubit concentrated using the dsDNA HS-assay (Invitrogen™ Qubit™ 3 Fluorometer, Waltham, MA), before being composited with other barcoded samples at 200 to 240 ng of DNA mass per sample. Each pool of approximately 150 samples were then purified using a QIAquick PCR Purification Kit (Qiagen, Valencia, CA) before being submitted for sequencing. Sequencing was conducted at the

Fralin Life Sciences Institute Genomic Sequencing Center at Virginia Tech on an Illumina MiSeq using a 250 cycle paired end protocol. Field blanks, trip blanks, filtration blanks, and DNA extraction blanks were included on each lane of sequencing and pooled using a maximum volume protocol instead of by mass.

Reads were processed using the QIIME2 pipeline [54] and annotated using the Greengenes [55] database (May 2013 release). Libraries were demultiplexed, trimmed to remove low quality reads, and then processed on a per lane basis using DADA2 for merging of paired end reads, denoising, chimera removal, and error correction. Generated amplicon sequence variants (ASVs) from DADA2 processing were merged from each lane of sequencing, clustered at 99%, and classified using the Greengenes database. Classified OTUs were binned if they contained the same full classification and nontarget sequences (e.g., sequences related to mitochondria and chloroplast) were removed. OTUs were then exported to be utilized for downstream data analysis.

Amplicons generated by the 515f/926r primer set were trimmed to the 806r primer site after DADA2 processing, to achieve an equivalent length for data processing as the 515f/806r primer set. To achieve this, the 515f/926r ASVs generated from DADA2 were exported from the QIIME2 pipeline and processed using the Cutadapt tool [56] in a 2-step process. First the Cutadapt tool was run using the 806r primer with default settings, trimming ASVs from the 515f/926r set down to the 515f/806r length. To account for potential base mismatches over the 20 bp primer produced during either the DADA2 error correction or sequencing base calls, the remaining untrimmed ASVs were run through the Cutadapt tool again with a max error rate (MER) set to 0.2, up from 0.1 in the default parameters. After trimming, the ASVs were imported back into QIIME2 and merged with the untrimmed reads from the 515f/806r primers before continuing with the pipeline outlined above. Details related to this process can be found in SI Figure A.1-G.2 and SI Table G.1-G.4.

Data Analysis

Beta Diversity and Principal Coordinate Analysis

Beta-diversity metrics were generated in QIIME2 using the core-metrics and core-metrics-phylogenetic plugins. Unweighted UniFrac distance matrices were employed to directly compare phylogenetic relationships among sequencing reads generated from multiple primer sets. Rarefaction was completed at both 3,500 reads (min read number per sample; presented due inclusion of low DNA samples of interest) and 10,000 reads. These distance matrixes were imported into R version 4.0.2 (Rstudio Team, 2020) to generate principal coordinate analysis (PCoA) plots and visualized using the QIIME2r (Bisanz, 2018) and ggplot2 version 3.3.2 packages in R-Studio version 4.0.2 (Wickham, 2016). Correspondence analysis was run at the Phylum and Genus level to identified taxa associated with specific sample groupings using the Phyloseq and Vegan packages with p-value cutoffs of 0.05 and 0.01 for Phylum and Genus comparisons, respectively.

Core and Discriminatory Microbiome Analysis

Classified, condensed, and filtered sequences were imported into R version 4.0.2 using QIIME2r to identify core and discriminatory OTUs between differing water types. Both

weighted and unweighted analyses were conducted on OTUs using two methods, with subsequent analysis primarily focusing on unweighted data, as it better accounted for rare and unique taxa. The first method was run using an OTU based approach (presented in the Appendix G Subsection 1) while the second used a binned classification-based approach which allowed for condensing OTUs with the same taxonomic classification. The latter method was conducted on each classification level from phylum to genus, with phylum and genus level analysis presented.

Due to the comprehensive nature of this study, with a wide range of operational conditions and different levels of specificity associated with the OTU-based and binned classification-based analysis methods, the core bacterial microbiota was defined independently for each analysis method. At the OTU level, unweighted taxonomically-assigned OTUs were averaged together for each water type, calculating a frequency of detection percentage for each identified OTU. Since OTUs have a high degree of specificity, OTUs were considered core if their frequency of detection was >70% in all samples belonging to the category of interest. Discriminatory OTUs were defined by two criteria: 1) any OTU that was found to be core in one water use and 2) at a frequency of detection less than twenty percent in the comparative group of interest, or if the difference in frequency of detection was greater than fifty percent between the two compared water uses.

For the binned core and discriminatory analysis, OTUs binned in the same taxonomic group were further transformed into an unweighted value on a per sample basis. Samples were averaged together similarly for the OTU-based analysis allowing for a calculated frequency of detection across samples of a given category at each taxonomic classification that was used to identify core and discriminatory taxa. Since multiple OTUs were binned to a singular taxonomic assignment, core and discriminatory taxa were subsequently defined by stricter criteria. Core taxa where any taxonomic assignments that were detected in greater than 80% of all samples from that water use, while discriminatory taxa were required to meet two specific criteria: 1) any taxa that were found to be core in one water use and 2) at a frequency of detection less than 20% in the compared water use or if the difference in frequency of detection was greater than 60% between the two compared water uses. Further, both analyses were conducted using a modified sample set that removed atypical systems identified using the beta diversity analysis.

Detailed metadata can be found in the SI Tables G.5- G.8 and includes information related to climate, treatment technology employed, disinfection method and residual, and intended application.

Statistical Analysis

All statistical analyses were conducted in R version 4.0.2 using the vegan package [58]. Statistical differences between generated sequences, including assessing the impact of trimming, were tested using analysis of variance (ANOVA). Unweighted UNIFRAC distance matrices were tested using the analysis of similarities (ANOSIM). A significance cutoff of $p < 0.05$ was applied for all statistical tests. Full statistical analyses are available in SI Tables G.9 through G.13.

RESULTS AND DISCUSSION

Microbiomes Vary by Water Use Category

A clear distinction was found in the composition of bacterial microbiota as a function of category of intended water use. Remarkably, this was the case across a very broad range of U.S. water systems, regardless of the impact imparted via various treatment trains or distribution systems, climates, or regional differences. PCoA of unweighted UniFrac distances, grouped by water use, are presented in Figure 6.1 (POC) and Figure 6.2 (POU) with additional PCoA plots presented in the SI Figures G.3-G.19. All factors tested were significantly associated with the composition of the bacterial microbiota, except when POC versus POU was compared across all samples (SI Figure G.4). Distinct microbiomes were encountered as a function of intended water use when categorized hierarchically both when categorized broadly (e.g., potable, potable reuse, non-potable reuse) and narrowly (e.g., potable bottled water, potable POC, potable POU, IPR POC, DPR POU) (ANOSIM $p < 0.05$; r-stat = 0.45 and 0.42 respectively). Comparison by type of disinfectant residual and DNA extraction kit also resulted in a statistically significant difference, but with lower r-statistics when compared to water use characterizations (ANOSIM, r-stat = 0.28 and 0.21). Samples derived from both DNA extraction kits were found to be well distributed throughout all water use categories, without clustering (SI Figure G.7). Bacterial microbiota also significantly differed when compared by geographical descriptors, such as climate and regional location, however their r-statistics were much lower than other identified factors (ANOSIM, r-stat = 0.06 and 0.17 respectively; SI Figure G.5 and G.6).

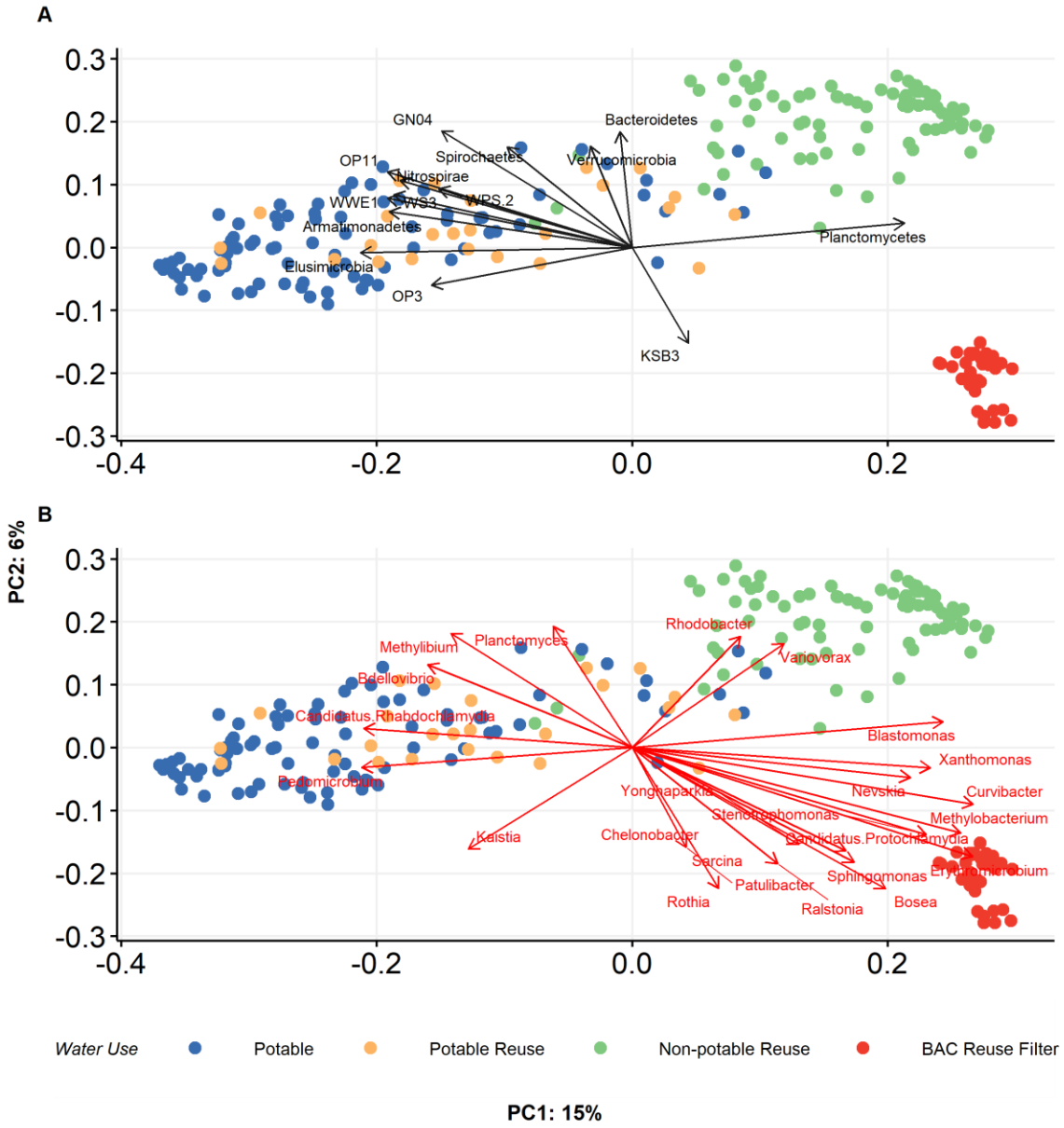


Figure 6.1: Unweighted UniFrac beta diversity plot for all bulk water samples at the POC, identified via their intended water use with overlayed correspondence analysis for (A) phylum with p-value cutoff of 0.05 and (B) genus with p-value cutoff of 0.01. Spectrum of tested water uses include potable (i.e., bottled water and conventional drinking water), potable reuse (e.g., direct potable reuse and indirect potable reuse), and non-potable reuse (i.e., wastewater secondary effluent with various degree of tertiary treatment) [37,39,51].

A clear visual difference in unweighted microbial community structure was identifiable at the POC as a function of intended water use category, even with a wide range of potable and non-potable water systems represented (ANOSIM, $p < 0.05$, r-stat = 0.75; Figure 6.2ab). Further, the bacterial composition of potable reuse water, including both IPR and DPR, overlapped substantially with that of conventional drinking water systems. The bacterial microbiota

associated with the BAC filter was a clear outlier, indicating that downstream processes strongly shape the signature of the finished potable reuse water.

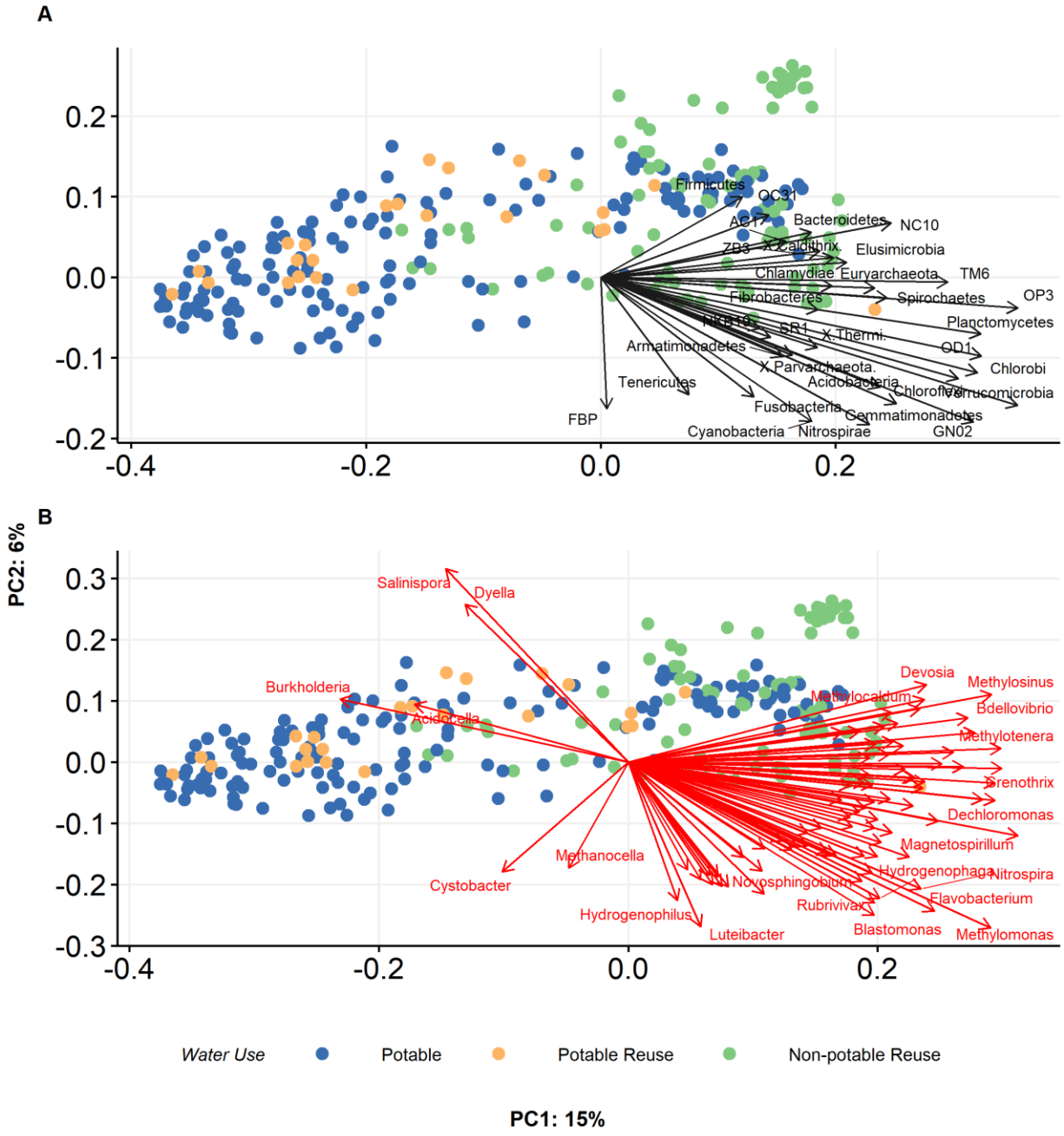


Figure 6.2: Unweighted UniFrac beta diversity plot for all bulk water samples at the POU, identified via their intended water use. Spectrum of tested water uses include potable (e.g. bottled water, and conventional drinking water), potable reuse (e.g. direct potable reuse, and indirect potable reuse), and non-potable reuse (e.g. wastewater secondary effluent with various degree of tertiary treatment) [37,39,51].

There was slightly less separation of bacterial microbiota composition according to intended water use category at the POU (ANOSIM, p-value < 0.05, r-stat = 0.34). Further, POC and POU were distinct specifically among potable reuse samples (ANOSIM, p-value < 0.05, r-stat = 0.16) and non-potable reuse samples (ANOSIM, p-value < 0.05, r-stat = 0.36), but not potable samples.

Various distinctions were also noted at the POC versus POU and as a function of various treatments. Changes in water quality and microbial composition from POC to POU are well-documented in potable water distribution systems [31,59–63]. Such shifts were also apparent in this study when each water use category was examined individually, especially the non-potable reuse and minimally-treated potable systems (i.e., those employing a singular process coupled with disinfection or disinfection alone, elaborated on in section 3.2). Few prior studies have assessed shifts in microbial communities in water reuse distribution systems [37,51,64]. Remarkably, the potable, potable reuse, and non-potable reuse bacterial microbiota compositions appeared to converge and become more similar at the POU relative to the POC (Figure 6.1 and Figure 6.2).

Potable System Microbiome Drivers

Variation encountered in the bacterial microbiota among potable water samples was associated with several factors, especially treatment train configuration (ANOSIM, p-value < 0.05, r-stat = 0.54) and the origin of initial source waters (e.g., surface water vs. groundwater) (ANOSIM, p-value < 0.05, r-stat = 0.60) (Figure 6.3ab). Additionally, microbial variation among these samples were explained to a lesser extent by climate (ANOSIM, p-value < 0.05, r-stat = 0.16) and region (ANOSIM, p-value < 0.05, r-stat = 0.42) (SI Figure G.8 and G.9). At the POC, correspondence analysis revealed that conventionally treated potable waters were associated with the phyla OP3 and Elusimicrobia (p-value < 0.05) and genus *Pedomicrobium* (p-value < 0.01) (Figure 6.1ab). At the POU, correspondence analysis revealed no phyla or genera associated with conventionally treated potable waters (Figure 6.2ab). Full lists of statistically significant taxa for all water uses can be found in the SI spreadsheet.

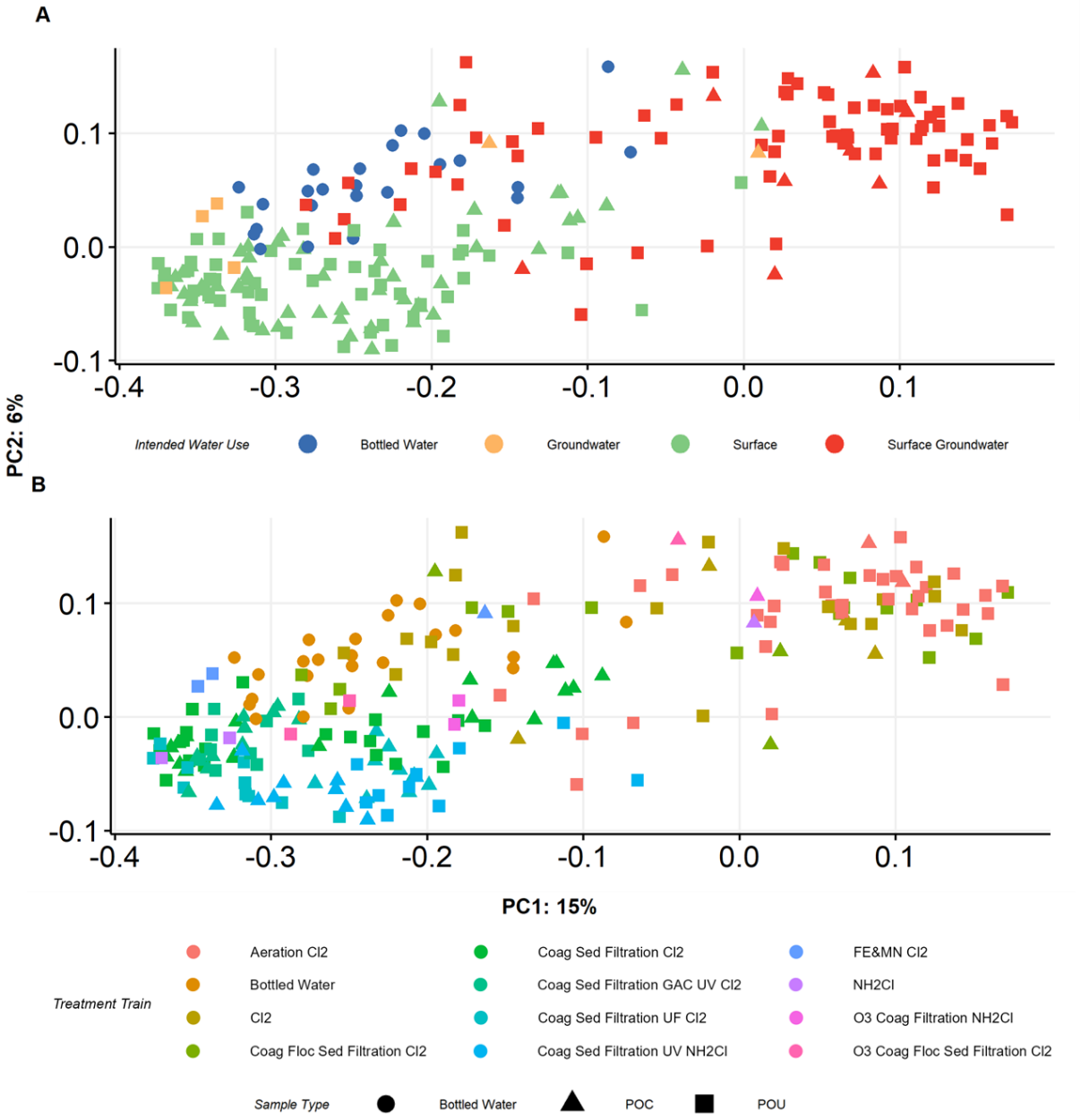


Figure 6.3: Unweighted UniFrac beta diversity plot for all bulk water samples intended for conventional potable use. Samples are classified by their (A) source water's origin and (B) respective treatment train and their representative sampling location at either the POC or POU.

Minimally-treated potable water treatment trains resulted in microbiomes that were more similar to non-potable reuse systems than potable water systems employing more comprehensive drinking water treatments (Figure 6.3ab). Specifically, systems utilizing only aeration and disinfection or relying only on disinfection resulted in microbial community structures that approached or overlapped non-potable reuse systems in bacterial composition, especially at the POU. Conversely, treatment trains that employed a combination of more conventional treatment processes; like coagulation, flocculation, sedimentation, filtration, membrane filtration, activated carbon filtration, and disinfection (e.g. chlorine, chloramine, and/or UV), were found to produce

a much more distinct bacterial microbiota, relative to non-potable reuse samples and minimally treated potable systems.

Source water was also a strong factor distinguishing among different potable water microbiomes. Notably, when overlaps between potable and non-potable water microbiomes were encountered, these tended to correspond to situations where the potable water treatments were minimal and potable water systems drew from a mix of groundwater and surface water sources. One possible explanation is that mixing of water sources expands the nutrient pool available to microbes, and minimal treatments accentuated the effects. However, minimal treatment in the absence of surface water did not result in a similar bacterial microbiota as those with mixed source waters. For example, iron and manganese removal coupled with disinfection resulted in a microbial community structure very similar to bottled water and in line with the conventionally-treated potable water systems. Still, even with stringent treatments, source water can have lingering effects. This was illustrated by the fact that samples with more extensive treatment trains (e.g. coagulation, flocculation, sedimentation, filtration, and chlorine disinfection) when fed a mixed source water still were found to develop microbial community structures more similar to minimally-treated potable systems and non-potable systems. This phenomenon was more apparent at the POU than the POC, suggesting that nutrient demand exerted in the distribution system led to population shifts. These findings provide new perspective on the need to tailor water treatment and distribution to the source water, especially for biologically unstable waters.

Although disinfection is intentionally employed to control microbial populations, it was interesting to find that the effects on the broader bacterial microbiota were much less pronounced than the other factors identified in this study. Disinfectant was a statistically significant factor, but wide dispersion was encountered in the PCoA plots for samples both from chlorine and chloramine systems (SI Figure G.10). Climate and regional locations were also found to be significant factors (SI Figure G.8 and G.9), e.g., Southwest versus Midwest U.S. regions. However, this could have been a result of the limited representation of climates and regions among other more dominant variables, such as treatment train configuration and/or source water origin.

Potable Reuse System Microbiome Drivers

Associations of examined factors with microbiomes for potable reuse systems were generally lower than those for conventional potable and non-potable reuse systems, with none resulting in r-statistics > 0.35 (SI Table G.13, with additional PCoA plots in the SI Figures G.15-G.18). Climate was one of the few factors that was associated with a significant difference in microbiome profiles among potable reuse samples (ANOSIM, p-value < 0.05, r-stat = 0.28), along with markers for various levels of treatment (Reuse Treatment – ANOSIM, p-value < 0.05, r-stat = 0.3; Post Blending Treatment - ANOSIM, p-value < 0.05, r-stat = 0.35). At the POC, correspondence analysis indicated association of the phyla Armatimonadetes, WWE1, WS3, WPS-2, Nitrospirae, and OP11 (p-value < 0.05) and genera *Candidatus Rhabdochlamydia* and *Bdellovibrio* (p-value <0.01) with potable reuse waters and the overlapped potable waters. Correspondence analysis did not reveal any phyla to be associated with potable reuse waters at the POU. At the genus level, *Burkholderia* and *Acidocella* (p-value < 0.01) were generally associated with the potable reuse and potable waters at the POU.

Even with substantial variety in treatment train configuration and corresponding bacterial microbiota, each potable reuse system produced a finished water with little to no overlap with non-potable reuse waters. This suggests that a wide range of potable reuse treatment options are capable of producing waters characterized by comparable potable benchmarks. This is an encouraging finding, supporting the concept that microbiome signatures have the potential to serve as a high resolution means to categorize reuse waters as potable or non-potable.

Non-potable Reuse System Microbiome Drivers

Bacterial microbiota in non-potable reuse systems varied as a function of stage of treatment (ANOSIM, $p < 0.05$, r-stat = 0.79; Figure 6.4). There was also a significant difference in the microbiota as a function of type of disinfection (ANOSIM, p-value < 0.05, r-stat = 0.42) and whether the system was undisinfected versus carrying a chloramine residual (ANOSIM, p-value < 0.05, r-stat = 0.34; SI Figure G.12). Notably, a subset of POU samples shifted towards a composition resembling that of the potable system POU samples, while others became more similar to the potable reuse biofilter effluent (SI Figure G.11). Across the wide range of bacterial microbiota observed across the distribution systems, geographical location (ANOSIM, p-value < 0.05, r-stat = 0.52) and climate (ANOSIM, p-value < 0.05, r-stat =0.38) were significant factors.

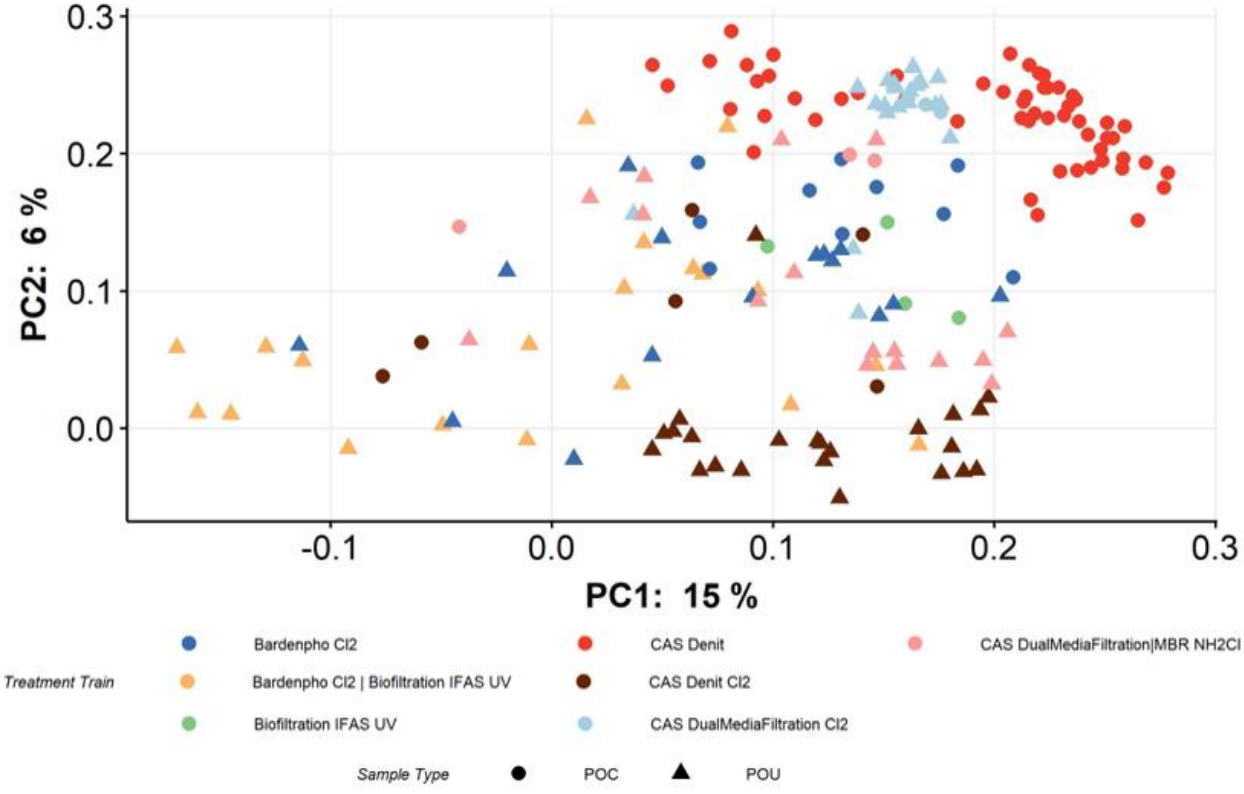


Figure 6.4: Unweighted UniFrac beta diversity plot for all bulk water samples intended for non-potable reuse. Samples are classified by their respective treatment train and their representative sampling location at either the POC or POU. Treatment train categories are indicated in Table 6.1.

At the POC, correspondence analysis revealed that the phyla GN04, Spirochaetes, Verrucomicrobia, and Bacteroidetes (p-value < 0.05) and genera *Methylibium,* and *Planctomyces* (p-value < 0.01) were associated with the interface between potable (both conventional and reuse) and non-potable reuse waters. At the POU, this interface was associated with the genera *Salinispora,* and *Dyella* (p-value < 0.01), while the more distant overlap between potable and non-potable reuse waters were associated with the phyla Firmicutes and OC31 (p-value < 0.05) and several genera, including: *Devosia, Methylocaldum, Methylosinus, Methylobacter, Lacibacter, and Perlucidibaca* (p-value < 0.01). Further, non-potable reuse samples that did not overlap with potable waters were found to be associated with the phylum Planctomycetes (p-value < 0.05) and genera *Rhodobacter, Variovorax,* and *Blastomonas* at the POC (p-value < 0.01).

At the POU, non-potable reuse waters were related to a wide range of phyla and genera. These included the phyla (p-value < 0.05): Bacteroidetes, Cyanobacteria, Planctomycetes, Acidobacteria, Chlamydiae, TM6, Verrucomicrobia, OD1, Spirochaetes, Chloroflexi, OP3, Chlorobi, GN02, and Fibrobacteres and the genera (p-value < 0.01): *Bdellovibrio, Dechloromonas, Desulfovibrio, Fluviicola, Paludibacter, Planctomyces,* and *Sediminibacterium* among others.

Similarly to potable systems, treatment train configuration and application of disinfection were the strongest factors explaining observed variation in the non-potable water use bacterial microbiota. Further, increased variance in microbial community composition at the POU appeared to be influenced by regional differences and changes in climate designations (SI Figure G.12- G.13). This is logical, given that less stringent treatments are typically applied in the production of non-potable reuse waters, which results in less biologically stable waters. These less biologically-stable waters are subsequently primed for other factors to shape the resultant bacterial microbiota [3,29,65–70], such as: distribution system residence times, age, and pipe material; consumption and inactivation of microbially suppressing disinfection residuals; composition of organics; inorganics; preceding treatments; water chemistry; and other geographical differences.

Phylum-Level Core and Discriminatory Analysis

Seventy-two unique phyla were identified across all water samples, with 5, 7, and 19 of these phyla found to be core to potable, potable reuse, and non-potable reuse systems, respectively, at some level (e.g., all samples combined, POC, and/or POU; Figure 6.5a). Proteobacteria were found to be core for all water uses and subcategories (e.g., all samples combined, POC, and POU) with a frequency of detection approaching 100% throughout. In addition to Proteobacteria, four phyla were found to be core in at least one subcategory of each water use, including: Actinobacteria (core to all except potable reuse POU), Bacteroidetes (core to all except potable reuse POU), Firmicutes (core to all except potable POU, potable reuse combined, and potable reuse POU), and Cyanobacteria (core to all except potable POC, potable POU, and potable reuse POU). These 5 ubiquitous phyla accounted for the entire core phyla associated with potable systems, which notably were found to be shared with both non-potable and potable reuse systems. Similarly, the remaining 2 core phyla associated with potable reuse waters were all core in non-potable reuse systems, which included Planctomycetes (core to potable reuse POC, and all non-potable reuse subcategorizations), and Chloroflexi (core to only potable reuse and non-potable reuse POC). The remaining 12 core phyla associated with only

non-potable reuse systems included: Acidobacteria, Chlamydiae, Verrucomicrobia, TM6, OD1, GN02, Chlorobi, Fusobacteria (only core at the POC), OP3, Spirochaetes, Tenericutes (only core at the POC), and Fibrobacteres (only core at the POC). In addition to the phylum Proteobacteria, Bacteroidetes and Planctomycetes were found at 100% frequency of detection in all non-potable reuse samples.

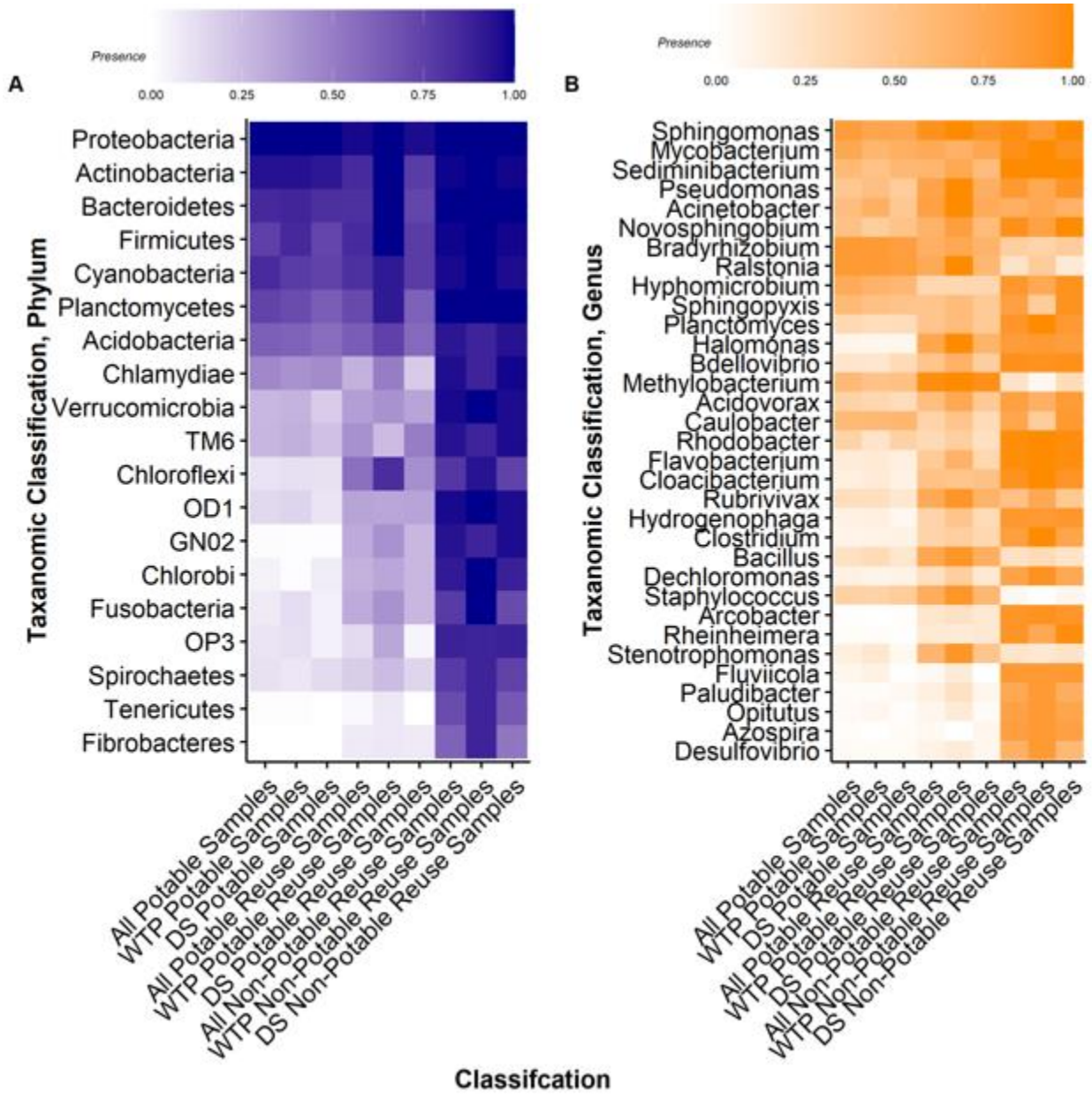


Figure 6.5: Binned unweighted core and discriminatory microbiome analysis conducted at the Phylum (A), and Genus (B) levels, grouped by water use and differentiations between the POC and POU. Average frequency of detection among all classified samples is presented. Core taxa where any taxonomic assignments that were present in greater than 80% of all samples from that water use while discriminatory taxa were required to meet two specific criteria: 1) any OTU that was found to be core in one water use and 2) at a prevalence less than 20% in the compared

water use or if the difference in prevalence was greater than 60% between the two compared water uses. Further, both analyses were conducted using a modified sample set that removed atypical systems identified using the beta diversity analysis. Associated core and discriminatory taxa are presented in Tables 2 and 3.

Several Phyla were identified that were characteristic of each water use category. There was overlap identified by core microbiome and correspondence analysis in 5 of 15 phyla at the POC and 13 of 15 phyla at the POU. Proteobacteria [71], Actinobacteria [72], and Bacteroidetes [73] were core to all water types, consistent with these Phyla tending to dominate aquatic environments and their wide range of metabolic functions. Cyanobacteria, known for association with algal blooms [74,75], were also core to all water types, underscoring the need to universally avoid conditions that trigger blooms. Firmicutes were core in non-potable systems and only directly after treatment in potable and potable reuse systems, which could be related to their ability to form spores and survive harsh environments imposed by disinfectants/advanced treatments. Planctomycetes, commonly found in aquatic and marine environments, was detected in 100% of non-potable reuse samples and 92% of potable reuse treated waters. Planctomycetes is well known for its important role in carbon and nitrogen cycling [76], with some species capable of anaerobic ammonium oxidation [7]. Higher frequency of detection relative to potable waters could be related to higher nutrient loads in reuse waters. Similarly, Acidobacteria was found to be core to only non-potable reuse waters, which may be related to its metabolic diversity, especially the ability to utilize nitrite and produce exopolysaccharide (EPS), a key component of biofilms [77].

Discriminatory bacterial taxa were readily identified among the various water uses, even at the Phylum level. Among water uses, more discriminatory phyla were found at the POC (11) than at the POU (7) or when evaluating the POC and POU samples together (7). Interestingly, the same seven discriminatory phyla were identified in the combined samples as the POU itself. Thus, further analysis focused on the discriminatory phyla detected at either the POC or POU. This analysis revealed that the phylum Chloroflexi, which was discriminatory between potable and reuse systems, was core in both potable reuse and non-potable reuse POC waters. However, Chlorflexi was not discriminatory throughout the distribution systems and overall the frequency of detection of most phyla within potable systems was very similar to potable reuse waters (Figure 6.5a).

Further evaluation revealed several phyla that discriminated between different water uses at both the POC and POU: Chlorobi, GN02, OD1, OP3, TM6, and Verrucomicrobia. Chlorobi and OD1 were both core to non-potable reuse water and uncommon in both potable and potable reuse systems at both the POC and POU. GN02, OP3, and Verrucomicrobia were found to be core to non-potable reuse systems and uncommon in both potable and potable reuse systems at the POU, while only uncommon in potable systems directly after treatment at the POC. TM6 was found to be core to non-potable reuse water at both the POC and POU, but uncommon in potable waters at the POU and uncommon in potable reuse waters directly after treatment.

At the POC, the phyla Fibrobacteres, Spirochaetes, and Tenericutes were found to be core following non-potable reuse treatments and uncommon in both potable and potable reuse waters, while Fusobacteria was core to non-potable reuse finished water and only uncommon in potable systems. At the POU, the phylum Chlamydiae was found to be discriminatory between non-

potable reuse (core) and potable reuse waters (uncommon) and was the only phyla found to be discriminatory at the POU and not at the POC.

When examining the phyla that discriminated potable (conventional and reuse) and non-potable reuse systems, most had apparent associations related to nutrient cycling, niche metabolic functions, or tendency to thrive in symbiotic relationships. OD1 is known to contain multiple versatile energy pathways, including nitrogen and fatty acid metabolism, while lacking biosynthetic capability [78–81]. This suggests the need for ectosymbiotic or parasitic relationships with other organisms, where non-potable reuse waters were likely to house relevant hosts. Chloroflexi are commonly found in treatment plants designed to remove phosphorus and nitrogen, with the ability to contribute to floc formation, degrade complex polymeric organic compounds, and sometimes cause sludge bulking [82,83]. OP3, which has been found in a full-scale partial-nitritation anammox reactors and is believed to facilitate nitrate reduction based on its carriage of *nar* [nitrate reductase], *norZ* [nitric oxide reductase], *and nrf* [pentaheme nitrite reductase] genes within their genome [84]. Chlorobi is comprised of anerobic photoautotrophic green sulfur bacteria, which are dominant in partial-nitritation anammox reactors and are believed to convert excess nitrate back into nitrite [84,85]. GN02 also is commonly abundant in partial-denitrification / anammox reactors [86,87].

Overall, despite the low taxonomic resolution inherent to Phylum-level analysis, it was remarkable that there were significant differences between the core and discriminatory microbiomes across the water categories. In particular, OD1, OP3, Chlorobi, Verrucomicrobia and GN02 exhibited high frequencies of detection in non-potable waters and extremely low frequencies of detection in both conventional potable and potable reuse systems. Strictly at the POC, the phyla Fibrobacteres and Tenericutes also appeared to be good candidates for differentiating potable and potable reuse waters from non-potable reuse waters. Chloroflexi strongly differentiated potable water systems from both potable and non-potable reuse water systems (Figure 6.5a and Table 6.2).

Table 6.2: Phylum based core and discriminatory analysis of taxa identified by correspondence analysis, reported as frequency of detection (decimal %)[1].

Phylum	Potable			Potable Reuse			Non-Potable Reuse			Correspondence Analysis, p<0.05	
	POC and POU	POC	POU	POC and POU	POC	POU	POC and POU	POC	POU	POC	POU
Proteobacteria	**1.00**	**1.00**	**1.00**	**0.97**	**1.00**	**0.96**	**1.00**	**1.00**	**1.00**		
Actinobacteria	**0.94**	**0.94**	**0.92**	**0.84**	**1.00**	0.77	**0.99**	**1.00**	**0.98**		
Bacteroidetes	**0.85**	**0.86**	**0.82**	**0.82**	**1.00**	0.73	**1.00**	**1.00**	**1.00**	*	*
Cyanobacteria	**0.84**	0.77	0.74	**0.82**	**0.92**	0.77	**0.97**	**1.00**	**0.96**		*
Firmicutes	0.76	**0.85**	0.73	**0.84**	**1.00**	0.77	**0.99**	**1.00**	**0.98**		*
Planctomycetes	0.74	0.70	0.63	0.71	**0.92**	0.62	**1.00**	**1.00**	**1.00**	*	*
Acidobacteria	0.62	0.61	0.57	0.63	0.75	0.58	**0.93**	**0.88**	**0.94**		*
Chlamydiae	0.45	0.41	0.44	0.29	0.50	0.19	**0.96**	**0.88**	**0.98**		*
TM6	0.27	0.30	0.23	0.42	0.25	0.50	**0.94**	**0.88**	**0.96**		*
Verrucomicrobia	0.27	0.29	0.19	0.37	0.42	0.35	**0.97**	**1.00**	**0.96**	*	*

OD1	0.14	0.16	0.10	0.34	0.33	0.35	**0.97**	**1.00**	**0.96**		*
Spirochaetes	0.11	0.09	0.13	0.18	0.25	0.15	0.78	**0.88**	0.75	*	*
Chloroflexi	0.10	0.11	0.11	0.55	**0.83**	0.42	0.79	**0.94**	0.75		*
OP3	0.10	0.11	0.06	0.13	0.33	0.04	**0.88**	**0.88**	**0.88**	*	*
Fusobacteria	0.08	0.13	0.06	0.32	0.42	0.27	0.78	**1.00**	0.71		*
Chlorobi	0.05	0.01	0.07	0.29	0.33	0.27	**0.91**	**1.00**	**0.88**		*
GN02	0.01	0.01	0.01	0.32	0.42	0.27	**0.94**	**0.88**	**0.96**		*
Tenericutes	0.01	0.01	0.00	0.03	0.08	0.00	0.70	**0.88**	0.65		*
Fibrobacteres	0.00	0.00	0.00	0.08	0.08	0.08	0.61	**0.88**	0.53		*

[1] Core taxa are bolded with discriminatory taxa are highlighted in red.

Genus Level Core and Discriminatory Analysis

At the genus level, 852 unique genera were identified with 3, 11, and 16 being core to potable, potable reuse, and non-potable reuse systems, respectively (Figure 6.5b, Table 6.3). Only the genus *Sphingomonas* was found to be core (in at least one subcategorization) to all water uses. Potable and potable reuse samples both commonly contained *Ralstonia* (core at all subdivisions, except potable reuse combined and at the POU only), while the potable reuse and non-potable reuse systems shared core genera of *Pseudomonas* (potable reuse POC, non-potable reuse combined, and non-potable reuse POU), *Novosphingobium* (potable reuse POC, non-potable reuse combined, non-potable reuse POC, and non-potable reuse POU), and *Halomonas* (potable reuse POC, non-potable reuse combined, non-potable reuse POC, and non-potable reuse POU). Potable and non-potable reuse water shared no core genera. Potable systems were characterized by one core genus, *Bradyrhizobium*, which was unique to all subcategorizations. The unique, core genera associated with potable reuse water included: *Acinetobacter* (except POU), *Halomonas* (POC only), *Methylobacterium*, *Rubrivivax* (POC only), *Bacillus* (POC only), *Staphylococcus* (POC only), and *Stenotrophomonas* (POC only). Additionally, the unique, core genera associated with non-potable reuse systems included: *Mycobacterium*, *Sediminibacterium*, *Hyphomicrobium* (except POC), *Sphingopyxis* (except POC), *Planctomyces*, *Halomonas*, *Bdellovibrio*, *Acidovorax* (except POC), *Caulobacter* (POU only), *Rhodobacter*, *Flavobacterium*, *Cloacibacterium*, and *Hydrogenophaga*. Of all core genera, only *Sediminibacterium* was found at 100% frequency of detection in all non-potable reuse samples.

Table 6.3: Genus based core and discriminatory analysis with reference to taxa identified by correspondence analysis, reported as frequency of detection (decimal %)[1].

Genus	Potable			Potable Reuse			Non-Potable Reuse			Correspondence Analysis, p<0.01	
	POC and POU	POC	POU	POC and POU	POC	POU	POC and POU	POC	POU	POC	POU
Bradyrhizobium	**0.87**	**0.87**	**0.83**	0.68	0.75	0.65	0.42	0.38	0.43		
Ralstonia	**0.87**	**0.87**	**0.84**	0.74	**1.00**	0.62	0.24	0.44	0.18	*	
Sphingomonas	**0.85**	0.78	0.77	**0.95**	**1.00**	**0.92**	**0.97**	**0.88**	**1.00**	*	*
Mycobacterium	0.76	0.63	0.69	0.74	0.67	0.77	**0.96**	**1.00**	**0.94**		*
Hyphomicrobium	0.73	0.66	0.61	0.34	0.33	0.35	**0.90**	0.75	**0.94**		*

Caulobacter	0.60	0.61	0.62	0.37	0.50	0.31	0.78	0.44	**0.88**		
Methylobacterium	0.60	0.52	0.53	**0.97**	**1.00**	**0.96**	0.24	0.06	0.29	*	
Sediminibacterium	0.60	0.53	0.61	0.63	0.75	0.58	**1.00**	**1.00**	**1.00**		*
Sphingopyxis	0.60	0.54	0.51	0.50	0.58	0.46	**0.82**	0.44	**0.94**		*
Acinetobacter	0.56	0.69	0.48	**0.82**	**1.00**	0.73	0.69	0.75	0.67		
Novosphingobium	0.53	0.43	0.51	0.66	**0.83**	0.58	**0.96**	**0.81**	**1.00**		*
Pseudomonas	0.48	0.55	0.44	0.79	**1.00**	0.69	**0.88**	0.75	**0.92**		
Staphylococcus	0.39	0.38	0.43	0.71	**0.92**	0.62	0.06	0.00	0.08		
Acidovorax	0.38	0.34	0.30	0.55	0.75	0.46	**0.82**	0.69	**0.86**		*
Rhodobacter	0.38	0.23	0.39	0.32	0.42	0.27	**0.99**	**1.00**	**0.98**	*	*
Planctomyces	0.33	0.29	0.30	0.50	0.58	0.46	**0.91**	**1.00**	**0.88**	*	*
Rubrivivax	0.27	0.28	0.20	0.74	**0.92**	0.65	0.54	0.75	0.47		*
Bacillus	0.26	0.31	0.21	0.76	**0.92**	0.69	0.25	0.31	0.24		
Bdellovibrio	0.22	0.21	0.31	0.50	0.67	0.42	**0.96**	**0.94**	**0.96**	*	*
Flavobacterium	0.17	0.18	0.13	0.45	0.67	0.35	**0.99**	**1.00**	**0.98**		*
Cloacibacterium	0.13	0.17	0.11	0.53	0.50	0.54	**0.94**	**1.00**	**0.92**		*
Stenotrophomonas	0.13	0.20	0.06	0.63	**0.92**	0.50	0.22	0.19	0.24		
Dechloromonas	0.11	0.09	0.10	0.26	0.42	0.19	0.79	**0.94**	0.75		*
Hydrogenophaga	0.11	0.13	0.07	0.37	0.50	0.31	**0.90**	**0.88**	**0.90**		*
Clostridium	0.10	0.10	0.11	0.37	0.42	0.35	**0.84**	**1.00**	0.78		
Halomonas	0.08	0.08	0.08	0.76	**1.00**	0.65	**0.85**	**0.88**	**0.84**		*
Fluviicola	0.08	0.11	0.01	0.05	0.17	0.00	**0.88**	**0.88**	**0.88**		*
Desulfovibrio	0.04	0.03	0.03	0.11	0.17	0.08	0.67	**0.88**	0.61		*
Opitutus	0.04	0.07	0.02	0.08	0.17	0.04	**0.81**	**0.88**	0.78		
Paludibacter	0.03	0.03	0.06	0.13	0.25	0.08	0.75	**0.88**	0.71		*
Azospira	0.02	0.00	0.03	0.05	0.00	0.08	**0.84**	**0.88**	**0.82**		
Rheinheimera	0.02	0.02	0.01	0.18	0.17	0.19	**0.93**	0.75	**0.98**		*
Arcobacter	0.01	0.00	0.01	0.21	0.25	0.19	**0.91**	**0.94**	**0.90**		*

[1] Core taxa are bolded with discriminatory taxa highlighted in red.

Core genera identified across the systems are also commonly encountered in aquatic environments. Core analysis and correspondence analysis overlapped in identification of 6 of 33 genera and 21 of 33 genera at the POC and POU, respectively. *Bradyrhizobium* and *Ralstonia*, were found to be core to potable systems and have been well documented in previous drinking water systems, including those producing ultra-pure water [88]. In such extreme oligotrophic environments, it is believed that *Bradyrhizobioum* plays an important role in the biodegradation of nitrogenous organic compounds. Various *Ralstonia* spp. are known to be denitrifiers and opportunities pathogens [89,90], and represents the only genus containing human pathogens that were core to potable systems.

When considering all three water uses, only a handful of core genera are known to contain human pathogenic members: the previously mentioned *Ralstonia*, *Mycobacteria* and *Arcobacter*. *Mycobacterium* spp. are known for slow growth rates, preference for biofilms, and high disinfection tolerance [91]. *Arcobacter spp.* are commonly found in wastewater effluents [92]

and have been identified as a key example of the need improved culture-independent measurements of effluent quality. Notably, *Arcobacter* was found to discriminate non-potable from all other waters, suggesting that it is well controlled by advanced treatment trains.

In terms of discriminatory genera, 14 unique relationships were found when both the POC- and POU-derived samples were combined and analyzed together (Figure 6.5b). When assessed independently, 12 were found at the POU and 20 at the POC. When assessing differences among genera core to non-potable systems and uncommon in potable and potable reuse systems, 3 genera (*Arcobacter*, *Azospira*, and *Fluviicola*) were found to be consistently discriminatory throughout all three comparisons, regardless of their origin relative to the POC or POU. The genera *Opitutus* and *Rheinheimera* displayed similar relationships among water use categories when all samples were combined, in addition to being discriminatory at only the POC and POU, respectively. Unlike these two cases, the genera *Paludibacter* and *Desulfovibrio* were only discriminatory directly after treatment, where a large reduction in frequency of detection in the distribution system caused the combined samples average to fail to meet the core genera criteria associated with the discriminatory analysis methodology.

A comparison among genera core to non-potable reuse systems and uncommon in only potable systems yielded three discriminatory genera (*Bdellovibrio*, *Cloacibacterium*, and *Hydrogenophaga*) being consistently discriminatory, regardless of the differentiations between the POC and POU. The genus *Halomonas* would also fall into this category, if not for it being identified as a core genus directly after potable reuse treatment. Similarly, *Flavobacterium* was also found to be consistently discriminatory between non-potable reuse and potable systems as it was found to be core in all non-potable reuse comparisons, uncommon in all potable comparisons, and uncommon in potable reuse systems only at the POU. The genus *Clostridium* was found at 100% frequency of detection in samples following non-potable reuse treatment and was also frequently detected at the POU (78%), resulting in it maintaining discriminatory classification across all samples. This was not the case for *Dechloromonas*, *Planctomyces*, and *Rhodobacter*, which were found to be discriminatory between potable and non-potable reuse waters only directly after treatment. Interestingly, *Rhodobacter* was also found to discriminate between non-potable reuse and potable reuse samples at the POU. When all samples were combined, *Rhodobacter* was discriminatory between non-potable reuse and uncommon in potable and potable reuse samples, mostly due to is extremely high frequency of detection in non-potable reuse water.

Beyond human pathogens, *Bdellovibrio* and *Flavobacterium* were found to be discriminatory between conventional potable waters and non-potable reuse waters. *Bdellovibrio* is parasitic to Gram negative bacteria [93] and was found in treatment trains relying on biological treatment (core to non-potable reuse and not uncommon in potable reuse water), where it was likely to readily find bacteria to prey upon. *Flavobacterium* is widely encountered in aquatic environments, especially wastewater treatment plants, and has the potential to cause disease in fish populations while also contributing to the degradation of complex organic matter [94–96]. Interestingly, *Cloacibacterium* shares the same family (Flavobacteriaceae) as *Flavobacterium* and has a similar ecological niche focused on complex organic matter degradation [95].

The main commonality among discriminatory genera appeared to be related to metabolic function and can either distinguish between non-potable and potable water, with or without the inclusion of potable reuse waters. This is logical, as some advanced treatment trains

implemented in potable reuse trains, especially those using biological treatment, may still utilize these niche metabolic pathways for complex or refractory organic removal. Examples of such genera (those not uncommon in potable reuse systems) included *Dechloromonas* and *Hydrogenophaga*, certain species of which are known to provide anerobic degradation of benzene, toluene, ethylbenzene, and xylene compounds [97] as well as perchlorate and nitrate reduction [98] and the degradation of various organic compounds (antibiotics included) while utilizing the oxidation of hydrogen as an energy source [99], respectively. The genera *Fluviicola*, *Desulfovibrio*, and *Azospira* provide similar distinctions between potable (both conventional and reuse) and non-potable waters, where *Fluviicola* has been abundantly found in an aerobic/anoxic deammonification process (DEMON), though not anticipated to provide nitrogen transformation [100], and *Desulfovibrio* which contain sulfur reducing bacteria [101]. *Azospira* has been found in nitrate contaminated aquifers and are believed to consist primarily of facultatively anaerobic denitrifiers also capable of perchlorate reduction [102,103] like the previously identified discriminatory genus *Dechloromonas.*

Potable reuse bacterial microbiota, although very similar to potable waters, were also associated with genera with unique metabolic roles. These included *Methylobacterium*, a genus able to metabolize formate, formaldehyde and methanol (a common carbon supplement utilized in tertiary treatment) [104], *Bacillus*, and *Stenotrophomonas*. The latter two were only found to be discriminatory directly after treatment, with *Bacillus* being able to secrete enzymes beneficial for bioremediation [105] and *Stenotrophomonas* being impactful for both nitrogen and sulfur cycling as well as bioremediation [106].

Of all discriminatory relationships identified, the genera *Arcobacter*, *Fluviicola*, and *Azospira* were identified as the most suitable candidates for distinguishing potable and non-potable water use, as they were all are consistently detected and abundant in non-potable reuse systems at the POC, persist at high frequency of detection throughout non-potable reuse distribution systems, and are uncommon at the POC and POU in both conventional potable systems and reuse systems intended for potable use. Strictly at the POC, the genera *Paludibacter* and *Desulfovibrio* are also good candidates, with *Dechloromonas* providing consistent differentiation between potable and non-potable reuse water, but not potable reuse water.

Common Features of Core and Discriminatory Microbiomes

From the above assessments, it was observed that most discriminatory relationships among water use categories were driven by genera that were core to non-potable reuse waters and the lack of detection of those same genera in potable and potable reuse systems. This type of relationship made up 37 out of the 46 total discriminatory relationships found at the genus level. Of the remaining 9 relationships, 7 were found to be discriminatory between genera core to potable reuse systems and uncommon in non-potable reuse or potable systems, while the remaining 2 were between core potable genera and uncommon non-potable reuse genera.

Methylobacterium was found to be core in potable reuse systems and consistently discriminatory between potable reuse and non-potable reuse systems, regardless of distinctions between the POC and POU. A similar relationship was found between *Staphylococcus* in non-potable and potable reuse systems, but only at the POC. *Bacillus* and *Stenotrophomonas* were found to be core in potable reuse water at the POC, but were uncommon in both non-potable reuse and potable POC samples, with no discriminatory relationships at the POU. *Rubrivivax*

was also found to be discriminatory directly after treatment between the potable reuse samples, where it was core, and potable samples, where it was not prevalent. This marks the only instance where a genus was found to be only core to potable reuse systems and uncommon in potable systems, a distinction only observed at the POC. When looking at discriminatory genera core to potable systems, only *Ralstonia* met the criteria and was found to be discriminatory at the POU and when POC and POU samples were combined. However, the frequency of detection of *Ralstonia* directly after treatment for non-potable reuse system samples (41%) precluded this relationship from being consistently experienced across both POC and POU comparisons.

Both genus and phylum based discriminatory analysis revealed that most core taxa were associated with non-potable samples from the POC. This shows that most core genera were related to the POC finished water, and that subsequent distribution led to increased microbial variation and shifts in microbial community composition that are likely related to localized factors. Overall, regardless of the level of taxonomic classification more discriminatory relationships were found directly after treatment than samples that were subjected to distribution, supporting the idea that distribution systems provide a microbial normalization between waters with different intended uses. Non-potable reuse systems were predominately found to have more core taxa and typically made up the core component of the discriminatory analysis, though other exemplary discriminatory relationships exist. Conventional potable systems and reuse systems intended for potable use also experienced extremely limited discriminatory components and shared many core taxa, while simultaneously being divergent from non-potable waters. This supports the overall hypothesis that intended water use, and the corresponding factors like employed treatment technologies, select for a distinct bacterial microbiota relative to its intended water use. This bacterial microbiota was identified, with candidate discriminatory taxa highlighted for future research, refinement, and validation via 16S rRNA gene amplicon sequencing while identifying a valuable application of next generation sequencing technologies for assessing and benchmarking finishing waters throughout the water and wastewater industries.

This work takes a substantial step towards identifying core and discriminatory microbiomes related to water use that can be utilized in assessing treatment train effectiveness from a broad microbial perspective, beyond examination of a handful of pathogens or fecal indicators. Such an approach could be of value when seeking to evaluate novel water treatment technologies and where on the spectrum of potable to non-potable signatures the resulting microbiomes lie.

ACKNOWLEDGEMENTS

We thank the Water Research Foundation Project 4961, the Hampton Roads Sanitation District, Spring Point Partners, LLC, and NSF NNCI Award 2025151for financial support of this research.

REFERENCES

1. Daims, H., Taylor, M. W. & Wagner, M. Wastewater treatment: a model system for microbial ecology. *Trends Biotechnol.* **24**, 483–489 (2006).

2. Bouwer, E. J. & Crowe, P. B. Biological Processes in Drinking Water Treatment. *J. Am. Water Works Assoc.* **80**, 82–93 (1988).

3. Prest, E. I., Hammes, F., van Loosdrecht, M. C. M. & Vrouwenvelder, J. S. Biological Stability of Drinking Water: Controlling Factors, Methods, and Challenges. *Front. Microbiol.* **0**, 45 (2016).

4. Johnson, D. R., Lee, T. K., Park, J., Fenner, K. & Helbling, D. E. The functional and taxonomic richness of wastewater treatment plant microbial communities are associated with each other and with ambient nitrogen and carbon availability. *Environ. Microbiol.* **17**, 4851–4860 (2015).

5. Zhu, G. *et al.* Biological Removal of Nitrogen from Wastewater. *Rev. Environ. Contam. Toxicol.* **192**, 159–195 (2008).

6. McIlroy, S. J. *et al.* Identification of active denitrifiers in full-scale nutrient removal wastewater treatment systems. *Environ. Microbiol.* **18**, 50–64 (2016).

7. Kuenen, J. G. Anammox bacteria: From discovery to application. *Nat. Rev. Microbiol.* **6**, 320–326 (2008).

8. Stokholm-Bjerregaard, M. *et al.* A critical assessment of the microorganisms proposed to be important to enhanced biological phosphorus removal in full-scale wastewater treatment systems. *Front. Microbiol.* **8**, 718 (2017).

9. Martín, H. G. *et al.* Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. *Nat. Biotechnol. 2006 2410* **24**, 1263–1269 (2006).

10. Gu, A. Z., Saunders, A., Neethling, J. B., Stensel, H. D. & Blackall, L. L. Functionally Relevant Microorganisms to Enhanced Biological Phosphorus Removal Performance at Full-Scale Wastewater Treatment Plants in the United States. *Water Environ. Res.* **80**, 688–698 (2008).

11. Nielsen, P. H., Saunders, A. M., Hansen, A. A., Larsen, P. & Nielsen, J. L. Microbial communities involved in enhanced biological phosphorus removal from wastewater — a model system in environmental biotechnology. *Curr. Opin. Biotechnol.* **23**, 452–459 (2012).

12. Mielczarek, A. T., Nguyen, H. T. T., Nielsen, J. L. & Nielsen, P. H. Population dynamics of bacteria involved in enhanced biological phosphorus removal in Danish wastewater treatment plants. *Water Res.* **47**, 1529–1544 (2013).

13. Helbling, D. E., Hollender, J., Kohler, H.-P. E., Singer, H. & Fenner, K. High-Throughput Identification of Microbial Transformation Products of Organic Micropollutants. *Environ. Sci. Technol.* **44**, 6621–6627 (2010).

14. Benner, J. *et al.* Is biological treatment a viable alternative for micropollutant removal in drinking water treatment processes? *Water Research* vol. 47 5955–5976 (2013).

15. Helbling, D. E., Johnson, D. R., Honti, M. & Fenner, K. Micropollutant Biotransformation Kinetics Associate with WWTP Process Parameters and Microbial Community Characteristics. *Environ. Sci. Technol.* **46**, 10579–10588 (2012).

16. Rizzo, L. *et al.* Consolidated vs new advanced treatment methods for the removal of contaminants of emerging concern from urban wastewater. *Sci. Total Environ.* **655**, 986–1008 (2019).

17. Krzeminski, P. *et al.* Performance of secondary wastewater treatment methods for the removal of contaminants of emerging concern implicated in crop uptake and antibiotic resistance spread: A review. *Sci. Total Environ.* **648**, 1052–1081 (2019).

18. Wolfe, R. L., Lieu, N. I., Izaguirre, G. & Means, E. G. Ammonia-oxidizing bacteria in a chloraminated distribution sytem: Seasonal occurrence, distribution, and disinfection resistance. *Appl. Environ. Microbiol.* **56**, 451–462 (1990).

19. Regan, J. M., Harrington, G. W., Baribeau, H., Leon, R. De & Noguera, D. R. Diversity of nitrifying bacteria in full-scale chloraminated distribution systems. *Water Res.* **37**, 197–205 (2003).

20. Wagner, D. & Chamberlain, A. H. L. Microbiologically influenced copper corrosion in potable water with emphasis on practical relevance. *Biodegrad. 1997 83* **8**, 177–187 (1997).

21. Zhu, Y. *et al.* Characterization of biofilm and corrosion of cast iron pipes in drinking water distribution system with UV/Cl2 disinfection. *Water Res.* **60**, 174–181 (2014).

22. Leclerc, H., Schwartzbrod, L. & Dei-Cas, E. Critical Reviews in Microbiology Microbial Agents Associated with Waterborne Diseases Microbial Agents Associated with Waterborne Diseases. *Crit. Rev. Microbiol.* **28**, 371–409 (2002).

23. Reynolds, K. A., Mena, K. D. & Gerba, C. P. Risk of Waterborne Illness Via Drinking Water in the United States. *Rev. Environ. Contam. Toxicol.* **192**, 117–158 (2008).

24. Paulson, C., Broley, W. & Stephens, L. *Blueprint for One Water*. (2017).

25. Paulson, C., Broley, W. & Stephens, L. Blueprint for One Water | The Water Research Foundation. *2017* (2017).

26. Mansfeldt, C. *et al.* Microbial residence time is a controlling parameter of the taxonomic composition and functional profile of microbial communities. *ISME J. 2019 136* **13**, 1589–1601 (2019).

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

28. Chao, Y., Mao, Y., Wang, Z. & Zhang, T. Diversity and functions of bacterial community in drinking water biofilms revealed by high-throughput sequencing. *Sci. Reports 2015 51* **5**, 1–13 (2015).

29. Pinto, A. J., Xi, C. & Raskin, L. Bacterial Community Structure in the Drinking Water Microbiome Is Governed by Filtration Processes. *Environ. Sci. Technol.* **46**, 8851–8859 (2012).

30. Gerrity, D. *et al.* Microbial community characterization of ozone-biofiltration systems in drinking water and potable reuse applications. *Water Res.* (2018) doi:10.1016/j.watres.2018.02.023.

31. Potgieter, S. *et al.* Long-term spatial and temporal microbial community dynamics in a large-scale drinking water distribution system with multiple disinfectant regimes. *Water Res.* **139**, 406–419 (2018).

32. Gomez-Alvarez, V., Revetta, R. P. & Domingo, J. W. S. Metagenomic analyses of drinking water receiving different disinfection treatments. *Appl. Environ. Microbiol.* **78**, 6095–6102 (2012).

33. Liu, T., Liu, S., Zheng, M., Chen, Q. & Ni, J. Performance Assessment of Full-Scale Wastewater Treatment Plants Based on Seasonal Variability of Microbial Communities via High-Throughput Sequencing. *PLoS One* **11**, e0152998 (2016).

34. Fang, D. *et al.* Microbial community structures and functions of wastewater treatment systems in plateau and cold regions. *Bioresour. Technol.* **249**, 684–693 (2018).

35. Leddy, M. B. *et al.* Characterization of Microbial Signatures From Advanced Treated Wastewater Biofilms. *J. Am. Water Works Assoc.* **109**, E503–E512 (2017).

36. Lin, Y., Li, D., Zeng, S. & He, M. Changes of microbial composition during wastewater reclamation and distribution systems revealed by high-throughput sequencing analyses. *Front. Environ. Sci. Eng. 2016 103* **10**, 539–547 (2016).

37. Garner, E. *et al.* Impact of blending for direct potable reuse on premise plumbing microbial ecology and regrowth of opportunistic pathogens and antibiotic resistant bacteria. *Water Res.* **151**, 75–86 (2019).

38. Stamps, B. W. *et al.* Characterization of the Microbiome at the World's Largest Potable Water Reuse Facility. *Front. Microbiol.* **0**, 2435 (2018).

39. Ji, P., Parks, J., Edwards, M. A. & Pruden, A. Impact of Water Chemistry, Pipe Material and Stagnation on the Building Plumbing Microbiome. *PLoS One* **10**, e0141087 (2015).

40. Shade, A. & Handelsman, J. Beyond the Venn diagram: The hunt for a core microbiome. *Environmental Microbiology* vol. 14 4–12 (2012).

41. Turnbaugh, P. J. *et al.* The Human Microbiome Project. *Nat. 2007 4497164* **449**, 804–810 (2007).

42. Methé, B. A. *et al.* A framework for human microbiome research. *Nature* **486**, 215–221 (2012).

43. Dewhirst, F. E. *et al.* The Human Oral Microbiome †. *J. Bacteriol.* **192**, 5002–5017 (2010).

44. Khoruts, A., Dicksved, J., Jansson, J. & Sadowsky, M. Changes in the composition of the human fecal microbiome after bacteriotherapy for recurrent Clostridium difficile-

associated diarrhea. *J. Clin. Gastroenterol.* **44**, 354–360 (2010).

45. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nat. 2012 4867402* **486**, 222–227 (2012).

46. Gill, S. R. *et al.* Metagenomic Analysis of the Human Distal Gut Microbiome. *Science* **312**, 1355 (2006).

47. Proctor, C. R. & Hammes, F. Drinking water microbiology-from measurement to management. *Curr. Opin. Biotechnol.* **33**, 87–94 (2015).

48. Cai, L., Ju, F. & Zhang, T. Tracking human sewage microbiome in a municipal wastewater treatment plant. *Appl. Microbiol. Biotechnol. 2013 987* **98**, 3317–3326 (2013).

49. Wu, L. *et al.* Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nat. Microbiol.* **4**, 1183–1195 (2019).

50. Garner, E. *et al.* Next generation sequencing approaches to evaluate water and wastewater quality. *Water Res.* **194**, 116907 (2021).

51. Garner, E. *et al.* Microbial Ecology and Water Chemistry Impact Regrowth of Opportunistic Pathogens in Full-Scale Reclaimed Water Distribution Systems. *Environ. Sci. Technol.* **52**, 9056–9068 (2018).

52. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J. 2012 68* **6**, 1621–1624 (2012).

53. Walters, W. *et al.* Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems* **1**, (2016).

54. Bolyen, E. *et al.* Author Correction: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2 (Nature Biotechnology, (2019), 37, 8, (852-857), 10.1038/s41587-019-0209-9). *Nature Biotechnology* vol. 37 1091 (2019).

55. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).

56. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).

57. Team, Rs. Citing RStudio – RStudio Support. *RStudio: Integrated Development for R. RStudio* https://support.rstudio.com/hc/en-us/articles/206212048-Citing-RStudio (2020).

58. Oksanen, J. *et al.* Package 'vegan' Title Community Ecology Package Version 2.5-7. (2020).

59. Perrin, Y., Bouchon, D., Delafont, V., Moulin, L. & Héchard, Y. Microbiome of drinking water: A full-scale spatio-temporal study to monitor water quality in the Paris distribution system. *Water Res.* **149**, 375–385 (2019).

60. El-Chakhtoura, J., Saikaly, P. E., van Loosdrecht, M. C. M. & Vrouwenvelder, J. S. Impact of Distribution and Network Flushing on the Drinking Water Microbiome. *Front. Microbiol.* **0**, 2205 (2018).

61. Lautenschlager, K. *et al.* A microbiology-based multi-parametric approach towards assessing biological stability in drinking water distribution networks. *Water Res.* **47**, 3015–3025 (2013).

62. Hwang, C., Ling, F., Andersen, G. L., LeChevallier, M. W. & Liu, W. T. Microbial community dynamics of an urban drinking water distribution system subjected to phases of chloramination and chlorination treatments. *Appl. Environ. Microbiol.* **78**, 7856–7865 (2012).

63. El-Chakhtoura, J. *et al.* Dynamics of bacterial communities before and after distribution in a full-scale drinking water network. *Water Res.* **74**, 180–190 (2015).

64. Kantor, R. S., Miller, S. E. & Nelson, K. L. The Water Microbiome Through a Pilot Scale Advanced Treatment Facility for Direct Potable Reuse. *Front. Microbiol.* **0**, 993 (2019).

65. Hull, N. M. *et al.* Drinking Water Microbiome Project: Is it Time? *Trends Microbiol.* **27**, 670–677 (2019).

66. Santos, Q. M. B. los *et al.* Emerging investigators series: microbial communities in full-scale drinking water distribution systems – a meta-analysis. *Environ. Sci. Water Res. Technol.* **2**, 631–644 (2016).

67. Pinto, A. J., Schroeder, J., Lunn, M., Sloan, W. & Raskin, L. Spatial-temporal survey and occupancy-abundance modeling to predict bacterial community dynamics in the drinking water microbiomez. *MBio* **5**, (2014).

68. Zhang, Y. & Liu, W.-T. The application of molecular tools to study the drinking water microbiome – Current understanding and future needs. *https://doi.org/10.1080/10643389.2019.1571351* **49**, 1188–1235 (2019).

69. Ling, F., Hwang, C., LeChevallier, M. W., Andersen, G. L. & Liu, W.-T. Core-satellite populations and seasonality of water meter biofilms in a metropolitan drinking water distribution system. *ISME J. 2016 103* **10**, 582–595 (2015).

70. Stanish, L. F. *et al.* Factors Influencing Bacterial Diversity and Community Composition in Municipal Drinking Waters in the Ohio River Basin, USA. *PLoS One* **11**, e0157966 (2016).

71. Vaz-Moreira, I., Nunes, O. C. & Manaia, C. M. Ubiquitous and persistent Proteobacteria and other Gram-negative bacteria in drinking water. *Sci. Total Environ.* **586**, 1141–1149 (2017).

72. Intro to acinobacter. *Nature* **196**, 1024–1024 (1962).

73. Gupta, R. S. The Phylogeny and Signature Sequences Characteristics of Fibrobacteres, Chlorobi, and Bacteroidetes. *http://dx.doi.org/10.1080/10408410490435133* **30**, 123–143

(2008).

74. Falconer, I. R. An Overview of Problems Caused by Toxic ( ) Blue-Green Algae Cyanobacteria in Drinking and Recreational Water. *Env. Toxicol* **14**, (1999).

75. He, X. *et al.* Toxic cyanobacteria and drinking water: Impacts, detection, and treatment. *Harmful Algae* **54**, 174–193 (2016).

76. Fuerst, J. A. & Sagulenko, E. Beyond the bacterium: planctomycetes challenge our concepts of microbial structure and function. *Nat. Rev. Microbiol. 2011 96* **9**, 403–413 (2011).

77. Kielak, A. M., Barreto, C. C., Kowalchuk, G. A., van Veen, J. A. & Kuramae, E. E. The Ecology of Acidobacteria: Moving beyond Genes and Genomes. *Front. Microbiol.* **0**, 744 (2016).

78. Nelson, W. C. & Stegen, J. C. The reduced genomes of Parcubacteria (OD1) contain signatures of a symbiotic lifestyle. *Front. Microbiol.* **0**, 713 (2015).

79. Castelle, C. J., Brown, C. T., Thomas, B. C., Williams, K. H. & Banfield, J. F. Unusual respiratory capacity and nitrogen metabolism in a Parcubacterium (OD1) of the Candidate Phyla Radiation. *Sci. Reports 2017 71* **7**, 1–12 (2017).

80. Tian, R. *et al.* Small and mighty: adaptation of superphylum Patescibacteria to groundwater environment drives their genome simplicity. *Microbiome 2020 81* **8**, 1–15 (2020).

81. Brown, C. T. *et al.* Unusual biology across a group comprising more than 15% of domain Bacteria. *Nat. 2015 5237559* **523**, 208–211 (2015).

82. Nierychlo, M. *et al.* The morphology and metabolic potential of the Chloroflexi in full-scale activated sludge wastewater treatment plants. *FEMS Microbiol. Ecol.* **95**, (2019).

83. Speirs, L. B. M., Rice, D. T. F., Petrovski, S. & Seviour, R. J. The Phylogeny, Biodiversity, and Ecology of the Chloroflexi in Activated Sludge. *Front. Microbiol.* **0**, 2015 (2019).

84. Speth, D. R., in 't Zandt, M. H., Guerrero-Cruz, S., Dutilh, B. E. & Jetten, M. S. M. Genome-based microbial ecology of anammox granules in a full-scale wastewater treatment system. *Nat. Commun. 2016 71* **7**, 1–10 (2016).

85. Yang, Z. *et al.* Algal biofilm-assisted microbial fuel cell to enhance domestic wastewater treatment: Nutrient, organics removal and bioenergy production. *Chem. Eng. J.* **332**, 277–285 (2018).

86. Cao, S. *et al.* Nitrite production from partial-denitrification process fed with low carbon/nitrogen (C/N) domestic wastewater: performance, kinetics and microbial community. *Chem. Eng. J.* **326**, 1186–1196 (2017).

87. Gilbert, E. M. *et al.* Low Temperature Partial Nitritation/Anammox in a Moving Bed

Biofilm Reactor Treating Low Strength Wastewater. *Environ. Sci. Technol.* **48**, 8784–8792 (2014).

88. Kulakov, L. A., McAlister, M. B., Ogden, K. L., Larkin, M. J. & O'Hanlon, J. F. Analysis of Bacteria Contaminating Ultrapure Water in Industrial Systems. *Appl. Environ. Microbiol.* **68**, 1548 (2002).

89. Ryan, M. P., Pembroke, J. T. & Adley, C. C. Differentiating the growing nosocomial infectious threats Ralstonia pickettii and Ralstonia insidiosa. *Eur. J. Clin. Microbiol. Infect. Dis.* **30**, 1245–1247 (2011).

90. Proctor, C. R., Edwards, M. A. & Pruden, A. Microbial composition of purified waters and implications for regrowth control in municipal water systems. *Environ. Sci. Water Res. Technol.* **1**, 882–892 (2015).

91. Fan, X. Y. *et al.* Functional genera, potential pathogens and predicted antibiotic resistance genes in 16 full-scale wastewater treatment plants treating different types of wastewater. *Bioresour. Technol.* **268**, 97–106 (2018).

92. Kristensen, J. M., Nierychlo, M., Albertsen, M. & Nielsen, P. H. Bacteria from the genus arcobacter are abundant in effluent from wastewater treatment plants. *Appl. Environ. Microbiol.* **86**, (2020).

93. Jurkevitch, E. The ecology of Bdellovibrio and like organisms in wastewater treatment plants. in *The Ecology of Predation at the Microscale* 37–64 (Springer, Cham, 2020). doi:10.1007/978-3-030-45599-6_2.

94. Bernardet, J.-F. & Bowman, J. P. The Genus Flavobacterium. in *The Prokaryotes* 481–531 (2006). doi:10.1007/0-387-30747-8_17.

95. Allen, T. D., Lawson, P. A., Collins, M. D., Falsen, E. & Tanner, R. S. Cloacibacterium normanense gen. nov., sp. nov., a novel bacterium in the family Flavobacteriaceae isolated from municipal wastewater. *Int. J. Syst. Evol. Microbiol.* **56**, 1311–1316 (2006).

96. Bernardet, J.-F., Nakagawa, Y. & Holmes, B. Proposed minimal standards for describing new taxa of the family Flavobacteriaceae and emended description of the family. *Int. J. Syst. Evol. Microbiol.* **52**, 1049–1070 (2002).

97. Coates, J. *et al.* Anaerobic benzene oxidation coupled to nitrate reduction in pure culture by two strains of Dechloromonas. *Nature* **411**, 1039–1043 (2001).

98. Nozawa-Inoue, M., Scow, K. M. & Rolston, D. E. Reduction of perchlorate and nitrate by microbial communities in vadose soil. *Appl. Environ. Microbiol.* **71**, 3928–3934 (2005).

99. Kämpfer, P. *et al.* Hydrogenophaga defluvii sp. nov. and Hydrogenophaga atypica sp. nov., isolated from activated sludge. *Int. J. Syst. Evol. Microbiol.* **55**, 341–344 (2005).

100. Gonzalez-Martinez, A. *et al.* Microbial community analysis of a full-scale DEMON bioreactor. *Bioprocess Biosyst. Eng. 2014 383* **38**, 499–508 (2014).

101. Voordouw, G. The genus Desulfovibrio: The centennial. *Applied and Environmental Microbiology* vol. 61 2813–2819 (1995).

102. Bellini, M. I., Gutiérrez, L., Tarlera, S. & Scavino, A. F. Isolation and functional analysis of denitrifiers in an aquifer with high potential for denitrification. *Syst. Appl. Microbiol.* **36**, 505–516 (2013).

103. Nam, J. H., Ventura, J. R. S., Yeom, I. T., Lee, Y. & Jahng, D. A novel perchlorate- and nitrate-reducing bacterium, Azospira sp. PMJ. *Appl. Microbiol. Biotechnol.* **100**, 6055–6068 (2016).

104. Gallego, V., García, M. T. & Ventosa, A. Methylobacterium hispanicum sp. nov. and Methylobacterium aquaticum sp. nov., isolated from drinking water. *Int. J. Syst. Evol. Microbiol.* **55**, 281–287 (2005).

105. Chen, K., Xie, S., Iglesia, E. & Bell, A. T. A PCR test to identify Bacillus subtilis and closely related species and its application to the monitoring of wastewater biotreatment. *Appl. Microbiol. Biotechnol.* **56**, 816–819 (2001).

106. Ryan, R. P. *et al.* The versatility and adaptation of bacteria from the genus Stenotrophomonas. *Nat. Rev. Microbiol. 2009 77* **7**, 514–525 (2009).

CHAPTER 7: CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

For expanded implementation and optimization of carbon-based water reuse treatment trains to occur, fundamental research needs to be conducted on the interplay between treatment processes, underlying microbial communities, their functions, and contaminant removal. While performance indicators provide insights into achievable effluent water quality, the optimization of biological systems requires higher resolution analysis to understand the microbial dynamics governing disinfection and biological removal so that areas susceptible to performance improvement can be identified. Further, high resolution molecular and organic analysis techniques provide insights into the ability of advanced treatment trains to sufficiently remove microbial and chemical contaminates of emerging concern (CEC) while identifying recalcitrant contaminants that require additional considerations or alterations to operation. This dissertation contributed heavily to this need through its comprehensive evaluation of the fate of biodegradable and non-biodegradable organic carbon fractions and microbial dynamics occurring within a carbon-based advanced treatment train intended for potable reuse applications. The research arcs investigated in this dissertation provide a jumping off point for more targeted applications of next generation sequencing (NGS) technologies to optimize treatment processes while expanding the role of NGS technologies as a monitoring tool.

Chapter 2 demonstrated a biodegradable dissolved organic carbon (BDOC) protocol adapted specifically to evaluating the biodegradable fraction of reuse water through treatment trains employing biofilm-based processes. Further, the value of deeper insights gained by BDOC analysis for achieving a critical understanding of the fate of particulate organic carbon (POC), BDOC, and non-biodegradable dissolved organic carbon (NBDOC) throughout an advanced treatment train, especially when compared to more conventional organic carbon profiling. Of particular interest were insights relayed to the interplay between BAC filtration and GAC contacting where inefficient BDOC removal during BAC treatment negatively impacts the GAC contactor's ability to removal NBDOC, shortening its lifespan and incurring higher capital costs. Discrepancies identified between BDOC removal potential and achieved removal can serve to guide attention towards optimization. Stringent removal of biodegradable carbon is especially beneficial for limiting undesirable downstream regrowth of microorganisms, e.g., in potable water distribution systems or in recharged aquifers. Ideally, a carbon-based AWT train should produce finished water that is as biologically stable as possible by minimizing effluent BDOC and its bioavailability, a distinction that conventional organic carbon profiling cannot assess. Further, benchmarking of BDOC results to operational conditions and water quality data proved useful in explaining BAC performance variability. Ultimately, these insights are critical when seeking to understand and ultimately optimize complex advanced water treatment trains.

In Chapter 3, shotgun metagenomic sequencing was utilized to provide a comprehensive assessment of the composition and fate of ARGs subjected to multiple stages of advanced water treatment. Specifically, this study profiled the complete resistome, identified core and discriminatory ARGs, determined the impact of individual treatment processes on concentrations of classified genes, assessed differences between relative and calculated absolute abundances, and differences between clinically relevant and bulk ARGs to better understand the complex nature of antibiotic resistance in potable reuse applications. The fate of ARGs and their cumulative resistome were heavily tied to treatment processes and the microbial communities

present. BAC filtration was identified as being a particularly important intermediate process as it commonly increased both relative and calculated absolute abundances. Ultimately, conclusions were highly dependent on measures of relative or calculated absolute abundances. Treatment processes, like ozonation and UV disinfection, were found to be effective at reducing total ARG concentrations while simultaneously increasing relative abundances. During the entire course of treatment, bulk ARG's experienced effective removal of calculated absolute abundances, while relative abundances typically increased, though both results were antibiotic resistance class dependent. Overall, treatment was found to be more effective at removing ARGs associated with clinical relevance, which were found to make up a much smaller fraction of the total resistome.

In the fourth chapter of this dissertation, functional metagenomic analysis was applied the assess the functional capacity of microorganisms present at different stages of treatment, identify the selection of discriminatory genes directly after treatment, and determined how operation impacted genes associated with specific functions of interest. In doing so, this work effectively combined high level functional characterizations and more specific functional analysis to elucidate the impact of treatment process, operational conditions, and water quality parameters on shaping functional characteristics. Assessment of genes associated with metabolic processes and nitrogen fixation were also found to provide unique insights into the biological functions occurring during BAC filtration and GAC contacting. Interestingly, it was found that while more bulk genes associated with metabolic process were found during BAC filtration, significantly more unique, degradation genes were found during GAC contacting. This, coupled with conclusions drawn from other sections, suggest that organic degradation during BAC filtration is heavily focused on readily bioavailable carbon generated during ozonation and lacks the selected pressures and operational conditions necessary to select for more complex, specialized metabolic process found during GAC contacting. Overall, functional metagenomics proved valuable in shedding light on the functional dynamics present within an advanced treatment train intended for potable reuse.

In the fifth chapter of this dissertation the efficacy of utilizing 16S rRNA gene amplicon sequencing to provide a comprehensive assessment of microbial communities was shown. It was found that there was a distinct microbiota associated with multiple stages of treatment, and that variation within those set communities were responsive to changing operation conditions, especially within BAF. Ozonation was found to exhibit the most volatility within the identified communities, with O_3:TOC ratio being related to those shifts. Within BAF, microbial communities were responsive to changing operation conditions, and exhibited extremely strong bifurcation with and without the addition of supplemental nutrients. Further, it was found that more taxa associated with specialized metabolic pathways were associated with GAC contacting than BAC filtration, suggesting that complex microbial degradation may occur more readily during GAC treatment than BAC filtration. Overall, this study improved the understanding of bacterial dynamics within a carbon-based advanced water treatment train which can be used to improve design and operation.

In sixth chapter microbial communities from multiple intended water uses were compared, including samples from conventional drinking water, bottled water, potable reuse water, and non-potable reuse water. Samples were collected from a wide range of system within each intended water use, and from a variety of climates and geographical locations. Ultimately, a comparison across all samples found that microbial communities clustered by their intended

water use, more-so than any other factor. Further testing identified that the variability of microbial communities within different water uses was best explained by degree of treatment, disinfection type, differences between the point of use and point of compliance, source water, and climate differences. Additionally, core and discriminatory taxa were identified between potable systems (both conventional and reuse) and non-potable systems at both the genus and phylum levels. The later was especially surprising given its low resolution. These findings are promising for the expansion of NGS technologies as an assessment and monitoring tool for advanced treatment trains by providing a way to benchmark effluent microbial communities beyond the presence or absence of pathogens.

Altogether, this research, through its implementation of next generation sequencing technologies and higher resolution organic profiling, advances the fundamental knowledge associated with microbial dynamics and organics removal present within and throughout multiple stages of a carbon-based water reuse treatment train. Though impactful, this research raises additional needs for further exploration of advanced treatment trains intended for water reuse.

In particular, continued research needs to be conducted on antibiotic resistance with the goal of better identifying relationships between ARG measurements (e.g., types and concentrations) and public health concerns, especially as they relate to risk. This type of research is extremely needed and equally difficult, as ARGs can transfer between organisms, potentially reaching a pathogenic host organism, and become selected for during exposure to chemical compounds. Further, research needs to be conducted to improve guidance on which measure of ARGs, relative or calculate absolute abundances, are more impactful to human health and which groups of ARG are of particular concern. This work is desperately needed in conjunction with continued research on mitigation strategies and technologies since this research highlights removal and selection mechanisms are not equally felt across all gene classifications. Having a better understanding of what is most important would allow for more targeted research on mitigation strategies and prevention of environmental dissemination.

Research needs to be conducted on the continued application of functional metagenomic sequencing which has been limitedly applied to drinking water systems, wastewater systems, and especially water reuse systems. This research has shown the promise of functional metagenomics to explore biological processes and the response of microorganisms to changing treatment and operation, highlighting room for optimization and continued development.

There are numerous aspects of the BAF study that are worthy of consideration for continued research revolving around optimization and further understanding of the fate of microbial CECs. Primarily, this research has shown the susceptibility of microbial populations present within the BAC filters and GAC contactors to respond to changing operational conditions, and ultimately fill metabolic niches. Further, the apparent distinction between microbial communities and their functional capacities present within both BAC and GAC treatments raises additional research questions about potential utilization in treatment processes. Therefore, continued research focused on inoculation and more targeted augmentation of microbial communities and functional process should prove beneficial in enhancing treatment performance. Additionally, the identified selection of ARGs within BAC filtrations raises numerous questions about the selective pressures for antibiotic resistance present in BAF

filtration and its implications for the creation of novel ARGs and their dissemination to the environment.

Currently, there also remains a need for more accurate predictive, bench-scale testing of ozone-BAF systems. Due to the nature of a BDOC analysis, it provides the potential for an expanded role in bench testing beyond just the characterization of organic contaminants present throughout treatment. Further research needs to be conducted to realize this expanded role and confirm the ability for BDOC analysis to predict BAF organic removal and assess differences in EBCTs.

Overall significant research needs remain for the continued optimization and intensification of carbon-based treatment systems intended for potable reuse, as such would directly improve the feasibility and sustainability of their implementation. NGS technologies have shown promise in assisting with this endeavor, but they themselves can benefit from work focused on the expansion of their role outside of strictly academic research and into more applied roles like water quality monitoring and/or surveillance.

APPENDIX A: SUPPLEMENTAL INFORMATION - BACKGROUND DATA ON PILOT DESIGN AND OPERATION

SI Table A.1 provides average measurements of general water quality parameters for the pilot influent and both GAC effluents to give context to the overall performance of the pilot system and associated influent water quality.

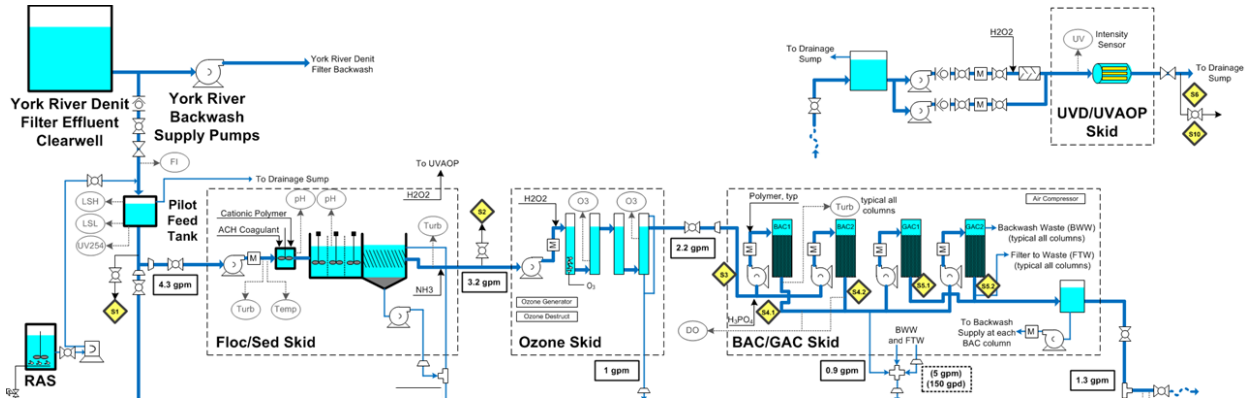
SI Table A.1: Measures of water quality of pilot influent and GAC effluent.

Parameter	Influent			GAC10 Effluent			GAC20 Effluent		
	Average	Sd	N	Average	Sd	N	Average	Sd	N
Temperature, Celsius	22.2	4.8	343	22.8	4.6	270	22.5	4.7	316
pH	7.3	0.2	343	-	-	-	-	-	-
Conductivity µS/cm	956.1	119.4	342	958.8	110.1	151	-	-	-
UVT %	72.6	3.4	329	90.4	4.0	385	92.9	3.5	305
TSS, minLOQ=1	5.9	2.8	92	-	-	-			
TN, minLOQ 0.5	2.4	1.8	72	2.2	1.3	68	2.3	1.4	71
Ammonia, minLOQ 0.2	0.5	0.4	61	0.2	0.0	38	0.4	1.0	41
TKN, minLOQ 0.5	1.6	0.8	73	0.6	0.1	68	0.7	0.7	71
NO3, minLOQ 0.2 mg/L	1.0	1.5	332	1.8	1.2	67	1.8	1.0	70
NO2, minLOQ 0.01	0.4	0.7	72	0.1	0.1	64	0.1	0.3	68
TP, minLOQ 0.2	0.4	0.2	69	0.1	0.1	68	0.1	0.1	71
Orthophosphate, minLOQ 0.01	0.27	0.35	304	0.04	0.03	66	0.02	0.02	42
TDS, minLOQ 1	584.0	62.9	56	-	-	-	-	-	-
Alkalinity, minLOQ 10	120.6	21.9	75	-	-	-	-	-	-
Magnesium, minLOQ 0.1	10.0	1.6	75	-	-	-	-	-	-
Hardness, minLOQ 0.1	191.6	17.8	75	-	-	-	-	-	-
Iron - Total, minLOQ 0.02	0.9	0.4	75	-	-	-	-	-	-
Manganese - Total, minLOQ 0.01	0.1	0.1	75	-	-	-	-	-	-
Iron - Dissolved, minLOQ 0.02	0.1	0.0	74	-	-	-	-	-	-
Manganese - Dissolved, minLOQ 0.01	0.1	0.1	74	-	-	-	-	-	-

TOC, minLOQ .1	7.9	1.4	75	3.9	0.9	418	3.1	0.9	416
DOC, minLOQ .1	7.0	1.2	75	-	-	-	-	-	-
COD, minLOQ 3	26.5	7.5	75	10.0	3.4	68	9.4	8.5	70
Turbidity, NTU	-	-	-	0.09	0.04	314	0.09	0.04	316

SI Figure A.1 and SI Figure A.2 provides context to the overall design and manufacture of the pilot system through a more detailed schematic and physical photo of the scale of the pilot system. More specifically, pilot influent waters from the YR-WWTP (conventional activated sludge plant with methanol fed Tetra^{TM} deep bed denitrification sand filters) denitrified secondary effluent were amended with return active sludge (ranging from 2 to 6 mg/L TSS) to more accurately represent a conventional activated sludge system's secondary effluent suspended solids concentrations. Coagulation-flocculation-sedimentation was completed via two stage coagulation (flow rate of 4.3 gpm, rapid mix HRT = 35 sec per stage, rapid mix G = 1000/sec, 25 mg/L of aluminum chlorohydrate, 0.75 mg/L cationic polymer) and three stage flocculation (setup in a tapered configuration; stage 1 G = 40/sec, stage 2 G = 20/sec, stage 3 G = 10/sec) while sedimentation was enhanced via a lamella plate settler.

Prior to ozonation, preformed monochloramine (3 mg Cl_2/L) was dosed into the settled water to reduce bromate formation. Ozonation itself (flow rate of 3.2 gpm) consisted of five columns operated in series where electrically generated ozone gas was diffused into the first column (diffused at the bottom via fine bubbles, traveling upwards) oppositive of the flow of water (feed at the top, traveling downward). Following ozonation, a chlorine quenching agent was typically added to prevent a significant chlorine load onto the BAC filters and GAC contactors.



SI Figure A.1: Detailed pilot schematic providing flow rates, potential chemical addition, and pumping configuration.

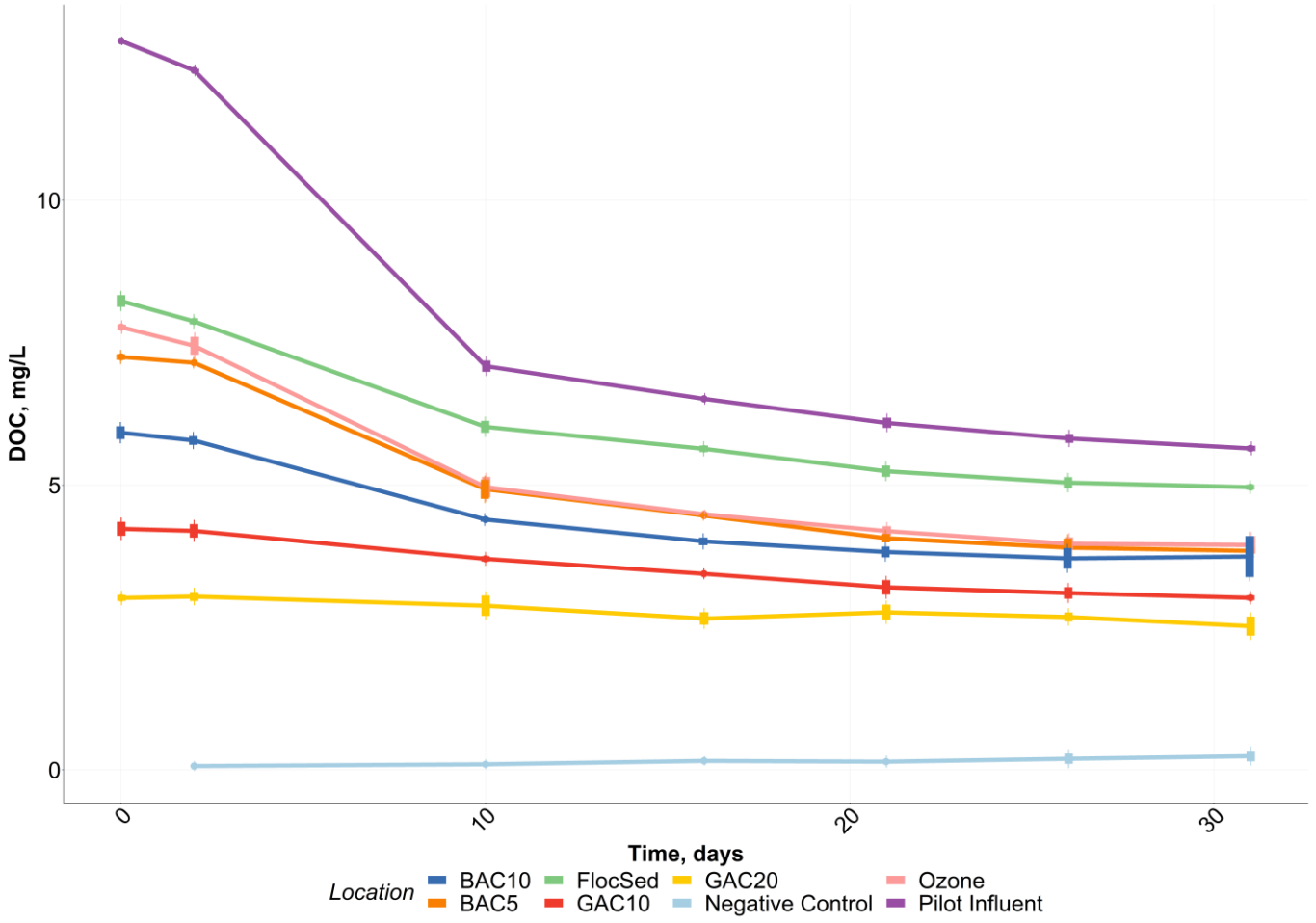
SI Figure A.2: Photograph of operational pilot plant with (1) coagulation, flocculation, and sedimentation skid, (2) ozone skid, and (3) BAC filter and GAC contactor skid.

BAC and GAC columns were operated in series with BAC5 feeding into GAC10, and BAC10 feeding into GAC20. More specific column design criteria for both the BAC and GAC systems can be found in SI Table A.2: BAC Filter and GAC Contactor Column Design Criteria.. Treated water leaving the GAC contactors was final irradiated with ultra-violet light with a target dose of 140 mJ/cm2.

SI Table A.2: BAC Filter and GAC Contactor Column Design Criteria.

Column	EBCT, min	Media Depth, ft	Flow Rate, gpm	Loading Rate, gpm/sf
BAC5	5.0	5.0	1.47	7.50
BAC10	10.0	6.0	0.73	3.70
GAC10	10.0	5.0	0.88	4.50
GAC20	20.0	6.0	0.44	2.60

APPENDIX B: SUPPLEMENTAL INFORMATION FOR CHAPTER 2



SI Figure B.1: Example degradation plot for sampling event SUMeOH-02 created during the BDOC analysis. Duration of the BDOC test is presented on the x-axis with each line indicating an average of triplicate reactors for each sample location with DOC measurements presented on the y-axis.

SI Table B.1: BDOC results broken up by sampling event, and sampling location

Sample Location	TOC Initial, mg/L							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	13.075	12.700	6.447	6.475	6.703	7.545	6.415	7.153
Floc Sed	8.485	9.378	5.122	5.195	5.108	5.968	4.383	5.460
Ozone	7.883	8.593	5.030	4.978	5.155	5.930	6.130	5.505
BAC5	7.298	7.893	3.905	3.643	4.525	4.803	6.143	5.273
GAC10	4.330	4.993	3.163	2.940	3.668	3.973	5.650	4.423
BAC10	5.913	6.377	3.628	3.365	4.073	4.385	5.173	4.948

GAC20	3.068	4.273	2.668	2.565	2.898	3.138	4.433	4.018
Sample Location	**DOC Initial, mg/L**							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	12.800	11.180	6.000	6.103	6.258	7.335	5.925	6.543
Floc Sed	8.230	7.545	4.987	5.128	4.878	5.558	4.213	4.953
Ozone	7.780	7.992	4.812	4.800	4.923	5.550	5.735	5.413
BAC5	7.250	7.437	4.007	3.425	4.188	4.750	5.948	5.143
GAC10	4.230	4.677	3.072	2.938	3.610	3.820	5.555	4.173
BAC10	5.920	6.697	3.593	3.315	4.070	4.388	4.948	4.690
GAC20	3.010	3.863	2.567	2.638	3.010	2.878	4.195	3.825
Sample Location	**Treatment Removal, mg/L**							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	-	-	-	-	-	-	-	-
Floc Sed	4.570	3.635	1.013	0.975	1.380	1.778	1.713	1.590
Ozone	0.450	-0.447	0.175	0.328	-0.045	0.007	-1.523	-0.460
BAC5	0.530	0.555	0.805	1.375	0.735	0.800	-0.212	0.270
GAC10	3.020	2.760	0.935	0.488	0.578	0.930	0.393	0.970
BAC10	1.860	1.295	1.218	1.485	0.852	1.163	0.788	0.723
GAC20	2.910	2.833	1.027	0.678	1.060	1.510	0.753	0.865
Sample Location	**Current Treatment Efficiency, %**							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	-	-	-	-	-	-	-	-
Floc Sed	35.70%	32.51%	16.89%	15.98%	22.05%	24.23%	28.90%	24.30%
Ozone	5.47%	-5.92%	3.51%	6.39%	-0.92%	0.13%	-36.14%	-9.29%
BAC5	6.81%	6.94%	16.73%	28.65%	14.93%	14.41%	-3.71%	4.99%
GAC10	41.66%	37.11%	23.33%	14.23%	13.79%	19.58%	6.60%	18.86%
BAC10	23.91%	16.20%	25.32%	30.94%	17.32%	20.95%	13.73%	13.35%
GAC20	49.16%	42.31%	28.57%	20.44%	26.04%	34.42%	15.21%	18.44%
Sample Location	**POC, mg/L**							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	0.275	1.520	0.447	0.373	0.445	0.210	0.490	0.610
Floc Sed	0.255	1.833	0.135	0.068	0.230	0.410	0.170	0.507
Ozone	0.103	0.602	0.218	0.178	0.233	0.380	0.395	0.093
BAC5	0.047	0.457	-0.102	0.218	0.338	0.053	0.195	0.130
GAC10	0.100	0.317	0.091	0.003	0.058	0.153	0.095	0.250
BAC10	-0.008	-0.320	0.035	0.050	0.002	-0.002	0.225	0.258
GAC20	0.058	0.410	0.102	-0.073	-0.113	0.260	0.238	0.193
Sample Location	**Change in POC, mg/L**							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	-	-	-	-	-	-	-	-
Floc Sed	-0.020	0.313	-0.312	-0.305	-0.215	0.200	-0.320	-0.103

Ozone	-0.153	-1.232	0.083	0.110	0.002	-0.030	0.225	-0.415
BAC5	-0.055	-0.145	-0.320	0.040	0.105	-0.327	-0.200	0.037
GAC10	0.053	-0.140	0.193	-0.215	-0.280	0.100	-0.100	0.120
BAC10	-0.110	-0.922	-0.183	-0.128	-0.230	-0.382	-0.170	0.165
GAC20	0.065	0.730	0.067	-0.123	-0.115	0.263	0.013	-0.065

Sample Location	**POC, %**							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	2.10%	11.97%	6.93%	5.75%	6.64%	2.78%	7.64%	8.53%
Floc Sed	3.01%	19.55%	2.64%	1.30%	4.50%	6.87%	3.88%	9.29%
Ozone	1.30%	7.00%	4.34%	3.57%	4.51%	6.41%	6.44%	1.68%
BAC5	0.65%	5.79%	-2.60%	5.97%	7.46%	1.09%	3.17%	2.47%
GAC10	2.31%	6.34%	2.88%	0.09%	1.57%	3.84%	1.68%	5.65%
BAC10	-0.13%	-5.02%	0.96%	1.49%	0.06%	-0.06%	4.35%	5.20%
GAC20	1.87%	9.59%	3.81%	-2.83%	-3.88%	8.29%	5.36%	4.79%

Sample Location	**Change in POC, %**							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	-	-	-	-	-	-	-	-
Floc Sed	0.90%	7.58%	-4.29%	-4.45%	-2.14%	4.09%	-3.76%	0.77%
Ozone	-1.70%	-12.55%	1.70%	2.27%	0.01%	-0.46%	2.56%	-7.61%
BAC5	-0.65%	-1.22%	-6.94%	2.41%	2.95%	-5.31%	-3.27%	0.79%
GAC10	1.66%	0.56%	5.48%	-5.89%	-5.89%	2.75%	-1.49%	3.19%
BAC10	-1.43%	-12.02%	-3.38%	-2.08%	-4.45%	-6.47%	-2.09%	3.52%
GAC20	2.00%	14.61%	2.85%	-4.31%	-3.94%	8.34%	1.01%	-0.41%

Sample Location	**BDOC, mg/L**							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	7.160	6.278	1.778	1.556	1.736	3.047	1.525	2.083
Floc Sed	3.270	3.043	1.198	0.609	-0.069	1.734	0.293	0.874
Ozone	3.830	4.818	2.047	2.273	1.631	2.095	2.098	2.261
BAC5	3.400	4.223	1.287	0.957	1.068	1.447	2.221	1.968
GAC10	1.210	1.732	0.724	0.728	0.760	0.757	2.493	1.291
BAC10	2.170	3.553	0.990	0.807	0.733	1.088	1.484	1.657
GAC20	0.490	1.315	0.443	0.633	0.442	0.288	1.408	1.053

Sample Location	**Change in BDOC, mg/L**							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	-	-	-	-	-	-	-	-
Floc Sed	-3.890	-3.235	-0.580	-0.947	-1.805	-1.313	-1.233	-1.208
Ozone	0.560	1.775	0.848	1.664	1.700	0.361	1.806	1.387
BAC5	-0.430	-0.595	-0.760	-1.317	-0.563	-0.648	0.123	-0.293
GAC10	-2.190	-2.492	-0.563	-0.229	-0.308	-0.690	0.272	-0.677
BAC10	-1.660	-1.265	-1.057	-1.467	-0.898	-1.008	-0.614	-0.604
GAC20	-1.680	-2.238	-0.547	-0.174	-0.292	-0.800	-0.076	-0.603

Sample Location	NBDOC (Final DOC), mg/L							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	5.640	4.902	4.222	4.547	4.522	4.288	4.400	4.460
Floc Sed	4.960	4.502	3.788	4.518	4.947	3.823	3.920	4.078
Ozone	3.950	3.173	2.765	2.527	3.292	3.455	3.637	3.152
BAC5	3.850	3.213	2.720	2.468	3.120	3.303	3.727	3.175
GAC10	3.020	2.945	2.348	2.210	2.850	3.063	3.063	2.882
BAC10	3.750	3.143	2.603	2.508	3.337	3.300	3.463	3.033
GAC20	2.520	2.548	2.123	2.005	2.568	2.590	2.787	2.772

Sample Location	Change in NBDOC, mg/L							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	-	-	-	-	-	-	-	-
Floc Sed	-0.680	-0.400	-0.433	-0.028	0.425	-0.465	-0.480	-0.382
Ozone	-1.010	-1.328	-1.023	-1.992	-1.655	-0.368	-0.283	-0.927
BAC5	-0.100	0.040	-0.045	-0.058	-0.172	-0.152	0.090	0.023
GAC10	-0.830	-0.268	-0.372	-0.258	-0.270	-0.240	-0.664	-0.293
BAC10	-0.200	-0.030	-0.162	-0.018	0.045	-0.155	-0.173	-0.118
GAC20	-1.230	-0.595	-0.480	-0.503	-0.768	-0.710	-0.677	-0.262

Sample Location	BDOC Fraction, %							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	55.94%	56.16%	29.64%	25.50%	27.74%	41.54%	25.74%	31.83%
Floc Sed	39.73%	40.34%	24.03%	11.88%	-1.42%	31.20%	6.94%	17.65%
Ozone	49.23%	60.29%	42.54%	47.36%	33.13%	37.75%	36.59%	41.77%
BAC5	46.90%	56.79%	32.11%	27.93%	25.49%	30.46%	37.34%	38.26%
GAC10	28.61%	37.03%	23.57%	24.77%	21.05%	19.81%	44.87%	30.94%
BAC10	36.66%	53.06%	27.55%	24.33%	18.02%	24.79%	30.00%	35.32%
GAC20	16.28%	34.04%	17.27%	23.98%	14.67%	9.99%	33.57%	27.54%

Sample Location	Change in BDOC, %							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	-	-	-	-	-	-	-	-
Floc Sed	-16.20%	-15.82%	-5.61%	-13.61%	-29.16%	-10.33%	-18.79%	-14.18%
Ozone	9.50%	19.96%	18.50%	35.48%	34.55%	6.54%	29.64%	24.12%
BAC5	-2.33%	-3.50%	-10.42%	-19.43%	-7.64%	-7.29%	0.75%	-3.51%
GAC10	-18.29%	-19.76%	-8.55%	-3.17%	-4.44%	-10.65%	7.53%	-7.32%
BAC10	-12.57%	-7.23%	-14.98%	-23.03%	-15.11%	-12.96%	-6.59%	-6.45%
GAC20	-20.38%	-19.02%	-10.28%	-0.35%	-3.34%	-14.80%	3.57%	-7.79%

Sample Location	NBDOC Fraction, %							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	44.06%	43.84%	70.36%	74.50%	72.26%	58.46%	74.26%	68.17%
Floc Sed	60.27%	59.66%	75.97%	88.12%	101.42%	68.80%	93.06%	82.35%
Ozone	50.77%	39.71%	57.46%	52.64%	66.87%	62.25%	63.41%	58.23%

BAC5	53.10%	43.21%	67.89%	72.07%	74.51%	69.54%	62.66%	61.74%
GAC10	71.39%	62.97%	76.43%	75.23%	78.95%	80.19%	55.13%	69.06%
BAC10	63.34%	46.94%	72.45%	75.67%	81.98%	75.21%	70.00%	64.68%
GAC20	83.72%	65.96%	82.73%	76.02%	85.33%	90.01%	66.43%	72.46%
Sample Location	**Change in NBDOC %**							
	S2	S4	S5	S6	S7	S8	S9	S10
Pilot Inf	-	-	-	-	-	-	-	-
Floc Sed	16.20%	15.82%	5.61%	13.61%	29.16%	10.33%	18.79%	14.18%
Ozone	-9.50%	-19.96%	-18.50%	-35.48%	-34.55%	-6.54%	-29.64%	-24.12%
BAC5	2.33%	3.50%	10.42%	19.43%	7.64%	7.29%	-0.75%	3.51%
GAC10	18.29%	19.76%	8.55%	3.17%	4.44%	10.65%	-7.53%	7.32%
BAC10	12.57%	7.23%	14.98%	23.03%	15.11%	12.96%	6.59%	6.45%
GAC20	20.38%	19.02%	10.28%	0.35%	3.34%	14.80%	-3.57%	7.79%

SI Table B.2: Summery stats of BDOC data including average, standard deviation, and median measurements

Measure of Central Tendency, Average

Location	Treatment Removal, mg/L	Treatment Efficiency, %	POC, mg/L	POC, %	BDOC, mg/L	NBDOC, mg/L	BDOC, %	NBDOC, %
Pilot Inf	-	-	0.55	6.5%	3.15	4.62	36.8%	63.2%
Floc Sed	2.08	25.1%	0.45	6.4%	1.37	4.32	21.3%	78.7%
Ozone	-0.19	-4.6%	0.28	4.4%	2.63	3.24	43.6%	56.4%
BAC5	0.61	11.2%	0.17	3.0%	2.07	3.20	36.9%	63.1%
GAC10	1.26	21.9%	0.13	3.0%	1.21	2.80	28.8%	71.2%
BAC10	1.17	20.2%	0.03	0.9%	1.56	3.14	31.2%	68.8%
GAC20	1.45	29.3%	0.13	3.4%	0.76	2.49	22.2%	77.8%

Measure of Central Tendency, Average

Location	TOC, mg/L	DOC, mg/L	ΔPOC, mg/L	ΔPOC, %	ΔBDOC, mg/L	ΔNBDOC, mg/L	ΔBDOC, %	ΔNBDOC, %
Pilot Inf	8.31	776.8%	-	-	-	-	-	-
Floc Sed	6.14	568.6%	-0.10	-0.2%	-1.78	-0.31	-15.5%	15.5%

Ozone	6.15	587.5%	-0.18	-2.0%	1.26	-1.07	22.3%	-22.3%
BAC5	5.44	526.8%	-0.11	-1.4%	-0.56	-0.05	-6.7%	6.7%
GAC10	4.14	400.9%	-0.03	0.0%	-0.86	-0.40	-8.1%	8.1%
BAC10	4.73	470.3%	-0.25	-3.5%	-1.07	-0.10	-12.4%	12.4%
GAC20	3.38	324.8%	0.10	2.5%	-0.80	-0.65	-9.0%	9.0%

Standard Deviation

Location	Treatment Removal, mg/L	Treatment Efficiency, %	POC, mg/L	POC, %	BDOC, mg/L	NBDOC, mg/L	BDOC, %	NBDOC, %
Pilot Inf	-	-	0.413	0.031	2.270	0.459	0.129	0.129
Floc Sed	1.306	0.070	0.577	0.059	1.232	0.481	0.153	0.153
Ozone	0.631	0.138	0.173	0.022	1.095	0.456	0.086	0.086
BAC5	0.459	0.097	0.176	0.033	1.176	0.460	0.105	0.105
GAC10	1.032	0.119	0.103	0.021	0.633	0.331	0.086	0.086
BAC10	0.386	0.062	0.176	0.031	0.938	0.421	0.107	0.107
GAC20	0.911	0.119	0.176	0.048	0.436	0.282	0.090	0.090

Standard Deviation

Location	TOC, mg/L	DOC, mg/L	ΔPOC, mg/L	ΔPOC, %	ΔBDOC, mg/L	ΔNBDOC, mg/L	ΔBDOC, %	ΔNBDOC, %
Pilot Inf	2.851	2.679	-	-	-	-	-	-
Floc Sed	1.795	1.420	0.244	0.043	1.168	0.346	0.068	0.068
Ozone	1.366	1.290	0.468	0.054	0.586	0.585	0.108	0.108
BAC5	1.550	1.490	0.167	0.036	0.410	0.093	0.062	0.062
GAC10	0.908	0.856	0.173	0.042	0.969	0.223	0.087	0.087
BAC10	1.068	1.145	0.313	0.045	0.378	0.089	0.056	0.056

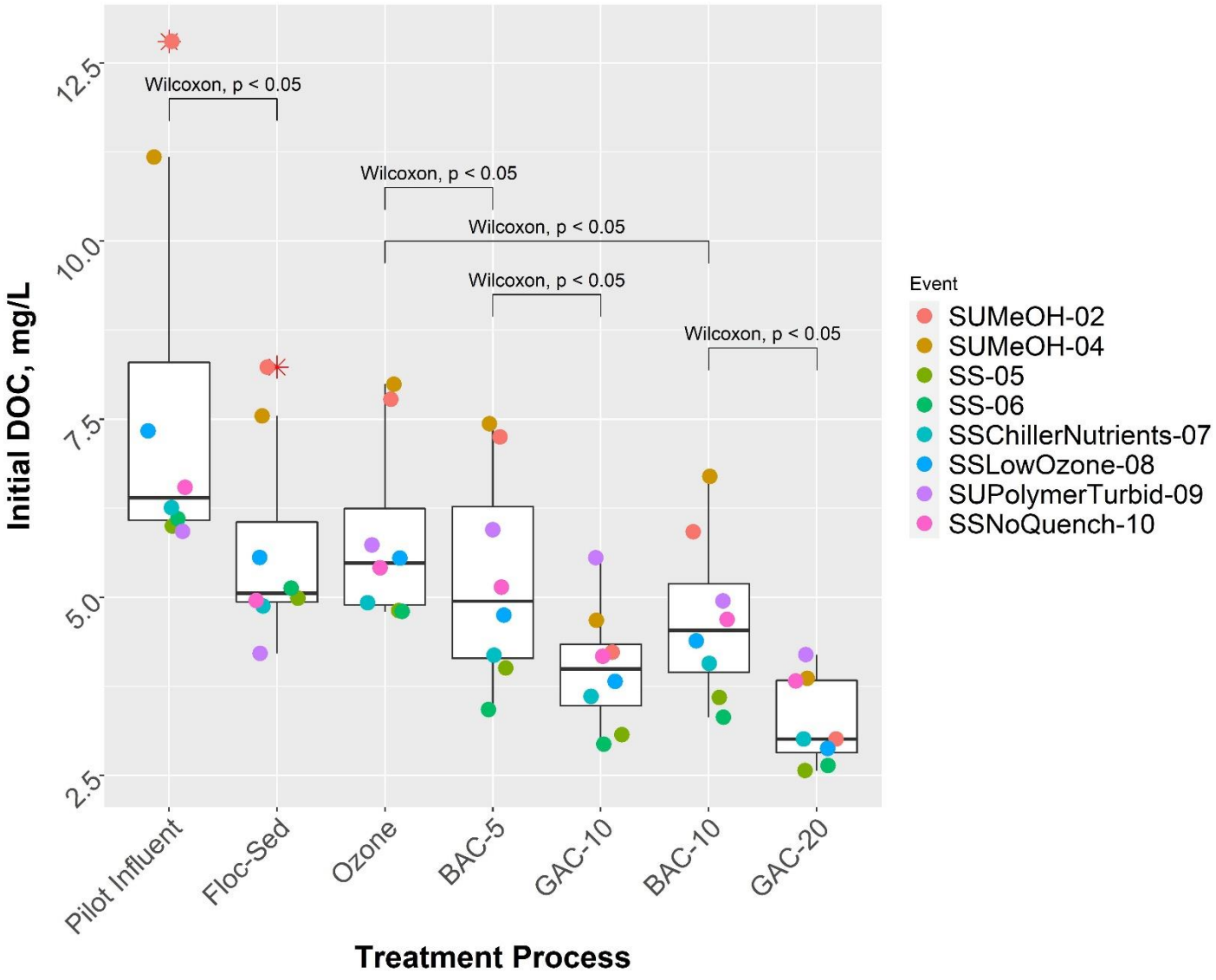
GAC20	0.744	0.621	0.282	0.063	0.767	0.283	0.087	0.087

Measure of Central Tendency, Median

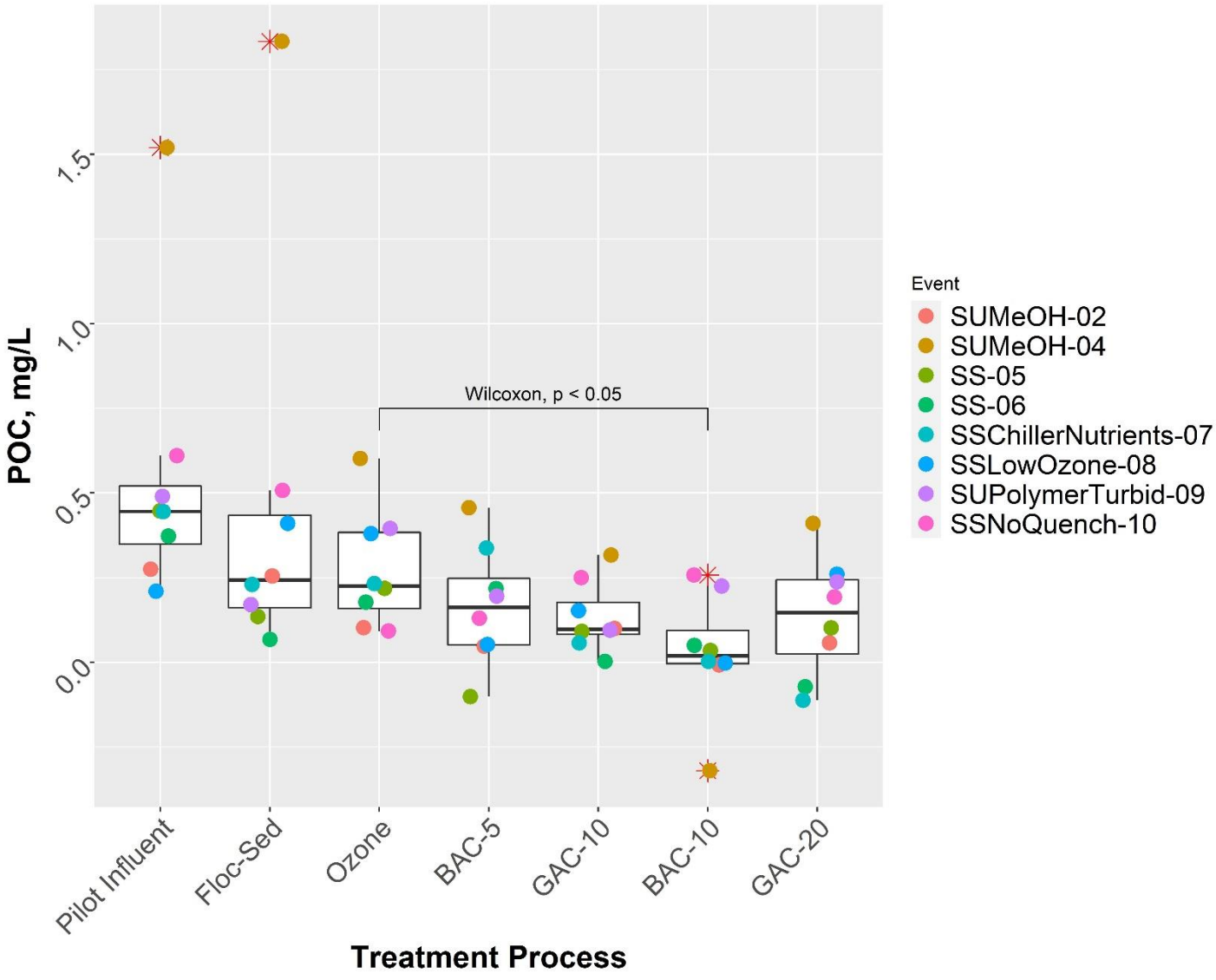
Location	Treatment Removal, mg/L	Treatment Efficiency, %	POC, mg/L	POC, %	BDOC, mg/L	NBDOC, mg/L	BDOC, %	NBDOC, %
Pilot Inf	-	-	0.45	6.8%	1.93	4.49	30.7%	69.3%
Floc Sed	1.65	24.3%	0.24	4.2%	1.04	4.29	20.8%	79.2%
Ozone	-0.02	-0.4%	0.23	4.4%	2.18	3.23	42.2%	57.8%
BAC5	0.65	10.7%	0.16	2.8%	1.71	3.19	34.7%	65.3%
GAC10	0.93	19.2%	0.10	2.6%	0.99	2.91	26.7%	73.3%
BAC10	1.19	19.1%	0.02	0.5%	1.29	3.22	28.8%	71.2%
GAC20	1.04	27.3%	0.15	4.3%	0.56	2.56	20.6%	79.4%

Measure of Central Tendency, Median

Location	TOC, mg/L	DOC, mg/L	ΔPOC, mg/L	ΔPOC, %	ΔBDOC, mg/L	ΔNBDOC, mg/L	ΔBDOC, %	ΔNBDOC, %
Pilot Inf	6.93	640.0%	-	-	-	-	-	-
Floc Sed	5.33	505.7%	-0.16	-0.7%	-1.27	-0.42	-15.0%	15.0%
Ozone	5.72	548.1%	-0.01	-0.2%	1.53	-1.02	22.0%	-22.0%
BAC5	5.04	494.6%	-0.10	-0.9%	-0.58	-0.05	-5.4%	5.4%
GAC10	4.15	399.6%	-0.02	1.1%	-0.62	-0.28	-7.9%	7.9%
BAC10	4.67	453.9%	-0.18	-2.7%	-1.03	-0.14	-12.8%	12.8%
GAC20	3.10	301.0%	0.04	1.5%	-0.58	-0.64	-9.0%	9.0%



SI Figure B.2: Initial DOC upon commencing BDOC analysis (effluent DOC for each treatment stage).



SI Figure B.3: POC concentration measured in the effluent of each treatment stage.

SI Table B.3: Statistical results for Wilcoxon signed rank test.

Data Measurement	p.value	Statistical Test	Comparison
BDOC.B10G20	0.0078125	Wilcoxon signed rank exact test	BAC10 and GAC20
BDOC.B5G10	0.0234375	Wilcoxon signed rank exact test	BAC5 and GAC10
BDOC.BB	0.0078125	Wilcoxon signed rank exact test	BAC5 and BAC10
BDOC.FO	0.0078125	Wilcoxon signed rank exact test	FlocSed and Ozone
BDOC.GG	0.0078125	Wilcoxon signed rank exact test	GAC10 and GAC20
BDOC.IF	0.0078125	Wilcoxon signed rank exact test	Influent and FlocSed
BDOC.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
BDOC.OB5	0.015625	Wilcoxon signed rank exact test	Ozone and BAC5
BDOCPERC.B10G20	0.0390625	Wilcoxon signed rank exact test	BAC10 and GAC20
BDOCPERC.B5G10	0.0546875	Wilcoxon signed rank exact test	BAC5 and GAC10
BDOCPERC.BB	0.0078125	Wilcoxon signed rank exact test	BAC5 and BAC10
BDOCPERC.FO	0.0078125	Wilcoxon signed rank exact test	FlocSed and Ozone
BDOCPERC.GG	0.0078125	Wilcoxon signed rank exact test	GAC10 and GAC20
BDOCPERC.IF	0.0078125	Wilcoxon signed rank exact test	Influent and FlocSed
BDOCPERC.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
BDOCPERC.OB5	0.015625	Wilcoxon signed rank exact test	Ozone and BAC5
CHANGEBDOC.B10G20	0.25	Wilcoxon signed rank exact test	BAC10 and GAC20
CHANGEBDOC.B5G10	0.7421875	Wilcoxon signed rank exact test	BAC5 and GAC10
CHANGEBDOC.BB	0.0078125	Wilcoxon signed rank exact test	BAC5 and BAC10
CHANGEBDOC.FO	0.0078125	Wilcoxon signed rank exact test	FlocSed and Ozone
CHANGEBDOC.GG	0.4609375	Wilcoxon signed rank exact test	GAC10 and GAC20
CHANGEBDOC.IF	0.0078125	Wilcoxon signed rank exact test	Influent and FlocSed
CHANGEBDOC.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
CHANGEBDOC.OB5	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC5
CHANGEBDOCPERC.B10G20	0.640625	Wilcoxon signed rank exact test	BAC10 and GAC20
CHANGEBDOCPERC.B5G10	0.84375	Wilcoxon signed rank exact test	BAC5 and GAC10
CHANGEBDOCPERC.BB	0.0078125	Wilcoxon signed rank exact test	BAC5 and BAC10
CHANGEBDOCPERC.FO	0.0078125	Wilcoxon signed rank exact test	FlocSed and Ozone
CHANGEBDOCPERC.GG	0.3828125	Wilcoxon signed rank exact test	GAC10 and GAC20
CHANGEBDOCPERC.IF	0.0078125	Wilcoxon signed rank exact test	Influent and FlocSed
CHANGEBDOCPERC.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
CHANGEBDOCPERC.OB5	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC5
CHANGENBDOC.B10G20	0.0078125	Wilcoxon signed rank exact test	BAC10 and GAC20
CHANGENBDOC.B5G10	0.0078125	Wilcoxon signed rank exact test	BAC5 and GAC10
CHANGENBDOC.BB	0.25	Wilcoxon signed rank exact test	BAC5 and BAC10

CHANGENBDOC.FO	0.0390625	Wilcoxon signed rank exact test	FlocSed and Ozone
CHANGENBDOC.GG	0.0234375	Wilcoxon signed rank exact test	GAC10 and GAC20
CHANGENBDOC.IF	0.0546875	Wilcoxon signed rank exact test	Influent and FlocSed
CHANGENBDOC.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
CHANGENBDOC.OB5	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC5
CHANGENBDOCPERC.B10G20	0.640625	Wilcoxon signed rank exact test	BAC10 and GAC20
CHANGENBDOCPERC.B5G10	0.84375	Wilcoxon signed rank exact test	BAC5 and GAC10
CHANGENBDOCPERC.BB	0.0078125	Wilcoxon signed rank exact test	BAC5 and BAC10
CHANGENBDOCPERC.FO	0.0078125	Wilcoxon signed rank exact test	FlocSed and Ozone
CHANGENBDOCPERC.GG	0.3828125	Wilcoxon signed rank exact test	GAC10 and GAC20
CHANGENBDOCPERC.IF	0.0078125	Wilcoxon signed rank exact test	Influent and FlocSed
CHANGENBDOCPERC.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
CHANGENBDOCPERC.OB5	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC5
CHANGEPOC.B10G20	0.078125	Wilcoxon signed rank exact test	BAC10 and GAC20
CHANGEPOC.B5G10	0.3828125	Wilcoxon signed rank exact test	BAC5 and GAC10
CHANGEPOC.BB	0.293029259	Wilcoxon signed rank exact test	BAC5 and BAC10
CHANGEPOC.FO	0.84375	Wilcoxon signed rank exact test	FlocSed and Ozone
CHANGEPOC.GG	0.3828125	Wilcoxon signed rank exact test	GAC10 and GAC20
CHANGEPOC.IF	0.3125	Wilcoxon signed rank exact test	Influent and FlocSed
CHANGEPOC.OB10	0.640625	Wilcoxon signed rank exact test	Ozone and BAC10
CHANGEPOC.OB5	0.84375	Wilcoxon signed rank exact test	Ozone and BAC5
CHANGEPOCPERC.B10G20	0.1484375	Wilcoxon signed rank exact test	BAC10 and GAC20
CHANGEPOCPERC.B5G10	0.546875	Wilcoxon signed rank exact test	BAC5 and GAC10
CHANGEPOCPERC.BB	0.4609375	Wilcoxon signed rank exact test	BAC5 and BAC10
CHANGEPOCPERC.FO	0.84375	Wilcoxon signed rank exact test	FlocSed and Ozone
CHANGEPOCPERC.GG	0.3828125	Wilcoxon signed rank exact test	GAC10 and GAC20
CHANGEPOCPERC.IF	0.84375	Wilcoxon signed rank exact test	Influent and FlocSed
CHANGEPOCPERC.OB10	0.3828125	Wilcoxon signed rank exact test	Ozone and BAC10
CHANGEPOCPERC.OB5	0.84375	Wilcoxon signed rank exact test	Ozone and BAC5
CURRENTTREATMENTEFFICIENCY.B10G20	0.078125	Wilcoxon signed rank exact test	BAC10 and GAC20
CURRENTTREATMENTEFFICIENCY.B5G10	0.1484375	Wilcoxon signed rank exact test	BAC5 and GAC10
CURRENTTREATMENTEFFICIENCY.BB	0.0078125	Wilcoxon signed rank exact test	BAC5 and BAC10
CURRENTTREATMENTEFFICIENCY.FO	0.0078125	Wilcoxon signed rank exact test	FlocSed and Ozone
CURRENTTREATMENTEFFICIENCY.GG	0.015625	Wilcoxon signed rank exact test	GAC10 and GAC20
CURRENTTREATMENTEFFICIENCY.IF	0.0078125	Wilcoxon signed rank exact test	Influent and FlocSed
CURRENTTREATMENTEFFICIENCY.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
CURRENTTREATMENTEFFICIENCY.OB5	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC5
DOC.B10G20	0.0078125	Wilcoxon signed rank exact test	BAC10 and GAC20
DOC.B5G10	0.0078125	Wilcoxon signed rank exact test	BAC5 and GAC10

DOC.BB	0.0078125	Wilcoxon signed rank exact test	BAC5 and BAC10
DOC.FO	0.640625	Wilcoxon signed rank exact test	FlocSed and Ozone
DOC.GG	0.0078125	Wilcoxon signed rank exact test	GAC10 and GAC20
DOC.IF	0.0078125	Wilcoxon signed rank exact test	Influent and FlocSed
DOC.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
DOC.OB5	0.015625	Wilcoxon signed rank exact test	Ozone and BAC5
NBDOC.B10G20	0.0078125	Wilcoxon signed rank exact test	BAC10 and GAC20
NBDOC.B5G10	0.0078125	Wilcoxon signed rank exact test	BAC5 and GAC10
NBDOC.BB	0.25	Wilcoxon signed rank exact test	BAC5 and BAC10
NBDOC.FO	0.0078125	Wilcoxon signed rank exact test	FlocSed and Ozone
NBDOC.GG	0.0078125	Wilcoxon signed rank exact test	GAC10 and GAC20
NBDOC.IF	0.0546875	Wilcoxon signed rank exact test	Influent and FlocSed
NBDOC.OB10	0.0390625	Wilcoxon signed rank exact test	Ozone and BAC10
NBDOC.OB5	0.1953125	Wilcoxon signed rank exact test	Ozone and BAC5
NBDOCPERC.B10G20	0.0390625	Wilcoxon signed rank exact test	BAC10 and GAC20
NBDOCPERC.B5G10	0.0546875	Wilcoxon signed rank exact test	BAC5 and GAC10
NBDOCPERC.BB	0.0078125	Wilcoxon signed rank exact test	BAC5 and BAC10
NBDOCPERC.FO	0.0078125	Wilcoxon signed rank exact test	FlocSed and Ozone
NBDOCPERC.GG	0.0078125	Wilcoxon signed rank exact test	GAC10 and GAC20
NBDOCPERC.IF	0.0078125	Wilcoxon signed rank exact test	Influent and FlocSed
NBDOCPERC.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
NBDOCPERC.OB5	0.015625	Wilcoxon signed rank exact test	Ozone and BAC5
POC.B10G20	0.574934085	Wilcoxon signed rank exact test	BAC10 and GAC20
POC.B5G10	0.574934085	Wilcoxon signed rank exact test	BAC5 and GAC10
POC.BB	0.293029259	Wilcoxon signed rank exact test	BAC5 and BAC10
POC.FO	0.640625	Wilcoxon signed rank exact test	FlocSed and Ozone
POC.GG	0.9453125	Wilcoxon signed rank exact test	GAC10 and GAC20
POC.IF	0.3125	Wilcoxon signed rank exact test	Influent and FlocSed
POC.OB10	0.0390625	Wilcoxon signed rank exact test	Ozone and BAC10
POC.OB5	0.1484375	Wilcoxon signed rank exact test	Ozone and BAC5
POCPERC.B10G20	0.4609375	Wilcoxon signed rank exact test	BAC10 and GAC20
POCPERC.B5G10	0.9453125	Wilcoxon signed rank exact test	BAC5 and GAC10
POCPERC.BB	0.4609375	Wilcoxon signed rank exact test	BAC5 and BAC10
POCPERC.FO	0.7421875	Wilcoxon signed rank exact test	FlocSed and Ozone
POCPERC.GG	0.7421875	Wilcoxon signed rank exact test	GAC10 and GAC20
POCPERC.IF	0.84375	Wilcoxon signed rank exact test	Influent and FlocSed
POCPERC.OB10	0.078125	Wilcoxon signed rank exact test	Ozone and BAC10
POCPERC.OB5	0.3828125	Wilcoxon signed rank exact test	Ozone and BAC5
TOC.B10G20	0.0078125	Wilcoxon signed rank exact test	BAC10 and GAC20

TOC.B5G10	0.0078125	Wilcoxon signed rank exact test	BAC5 and GAC10
TOC.BB	0.0078125	Wilcoxon signed rank exact test	BAC5 and BAC10
TOC.FO	0.546875	Wilcoxon signed rank exact test	FlocSed and Ozone
TOC.GG	0.0078125	Wilcoxon signed rank exact test	GAC10 and GAC20
TOC.IF	0.0078125	Wilcoxon signed rank exact test	Influent and FlocSed
TOC.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
TOC.OB5	0.015625	Wilcoxon signed rank exact test	Ozone and BAC5
TREATMENTREMOVAL.B10G20	0.3125	Wilcoxon signed rank exact test	BAC10 and GAC20
TREATMENTREMOVAL.B5G10	0.25	Wilcoxon signed rank exact test	BAC5 and GAC10
TREATMENTREMOVAL.BB	0.0078125	Wilcoxon signed rank exact test	BAC5 and BAC10
TREATMENTREMOVAL.FO	0.0078125	Wilcoxon signed rank exact test	FlocSed and Ozone
TREATMENTREMOVAL.GG	0.1484375	Wilcoxon signed rank exact test	GAC10 and GAC20
TREATMENTREMOVAL.IF	0.0078125	Wilcoxon signed rank exact test	Influent and FlocSed
TREATMENTREMOVAL.OB10	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC10
TREATMENTREMOVAL.OB5	0.0078125	Wilcoxon signed rank exact test	Ozone and BAC5

Subsection 1: Benchmarking of GAC Contacting to Bed Volumes

The following data provides a benchmark of GAC contactor BDOC samples to their respective bed volumes at the time of sampling. Raw data can be found in SI Table A1, with SI Figure A1 providing a treatment efficiency breakthrough curve for GAC10 and GAC20 benchmarked to their respective bed volumes. Additionally, treatment efficiencies for each BDOC sampling are overlayed for both contactors. From the breakthrough curve the transient, adsorption-based removal begins to bottom out at around 20,000 bed volumes, after which removal mechanisms begin to primarily consistent of biological removal associated with colonized microorganisms.

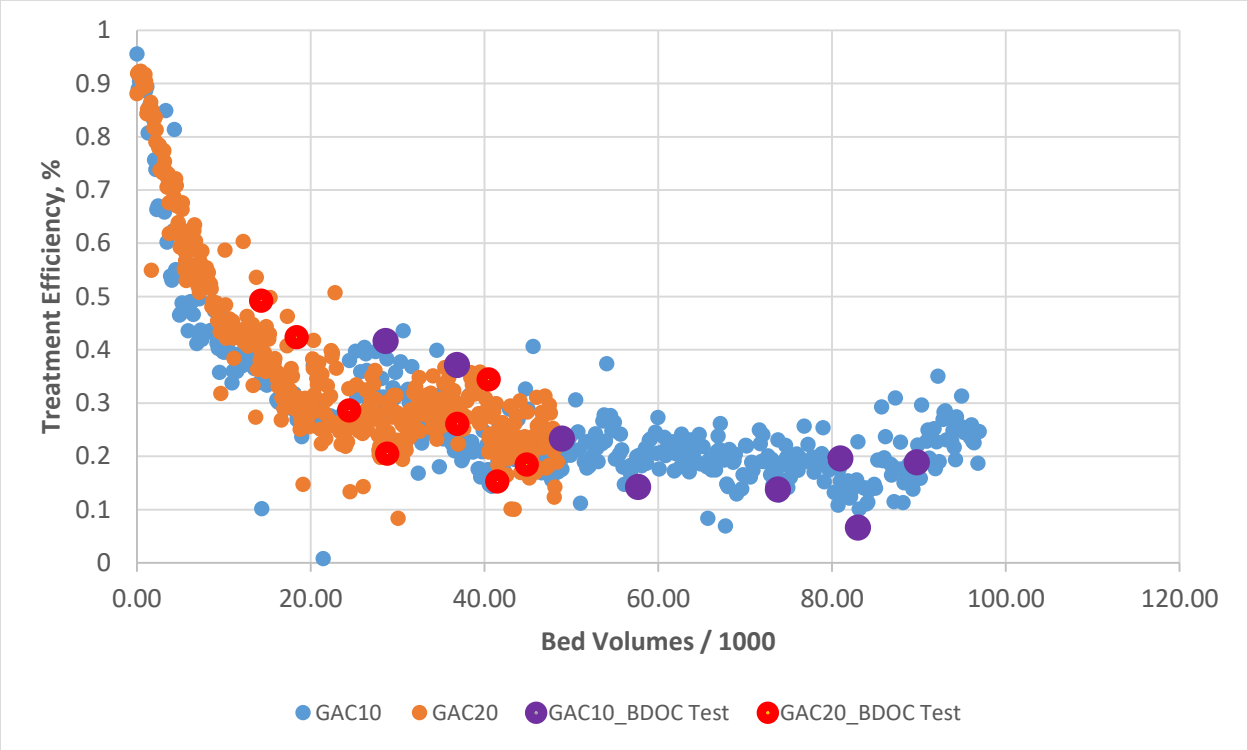
From this data, all but one sampling event occurred after 18,000 bed volumes (a period after adsorption-based removal mechanisms were dominant) with only a singular sample occurring before 15,000 bed volumes (GAC20, S02, 14308 bed volumes). This indicates that the provided GAC contacting analysis is primarily focused on GAC operation after the transient, adsorption dominated removal. This analysis therefore provides insights into prolonged GAC operation.

It is also noteworthy that the only sampling events below the identified 20k bed volumes threshold, SUMeOH-02 (S02) and SUMeOH-04 (S04), were also impacted by upstream methanol leaching and may have hold over effects resulting in elevated DOC fractions and removal. Therefore, elevated DOC removal (SI Figure A2), treatment efficiencies (SI Figure A3), and BDOC removal (SI Figure A4) for SUMeOH-02 and SUMeOH-04 are likely more a result of upset conditions than lower bed volumes. This is supported by the DOC removal, specifically the BDOC fraction, that occurred during SUMeOH-02 and SUMeOH-04 for GAC10 that was not seen in GAC20 at similar bed volumes when methanol leaching was not an issue. Further, NBDOC removal was seemingly unaffected after 15,000 bed volumes indicating that

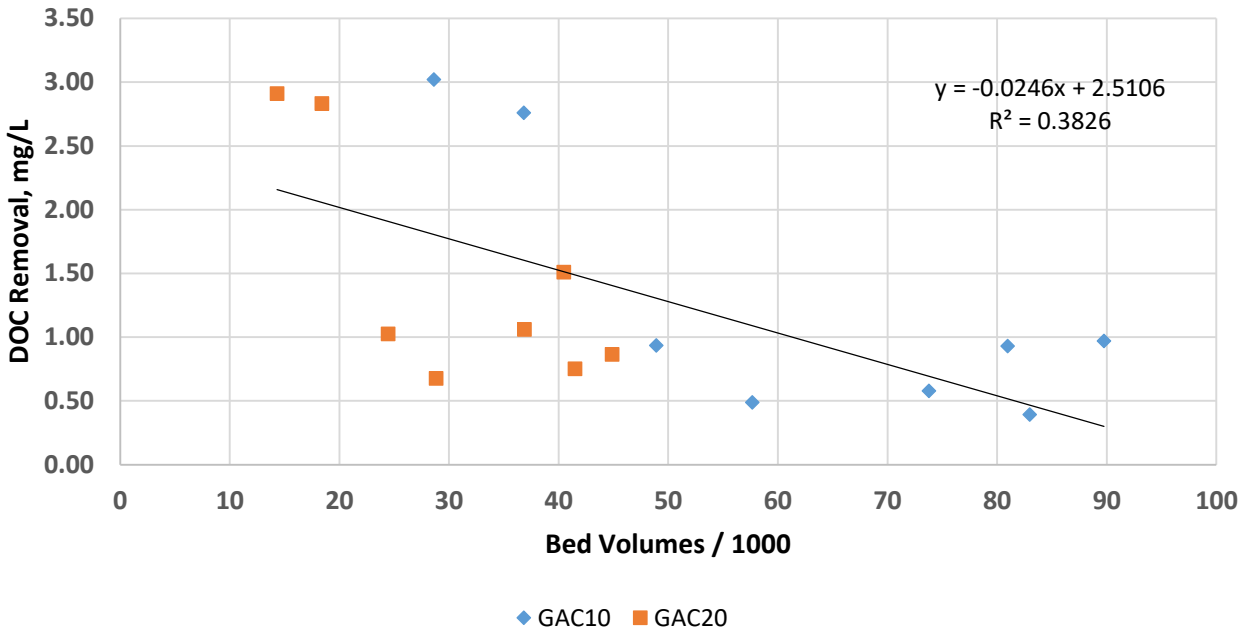
increase DOC removal was heavily influenced by BDOC. Otherwise, DOC and BDOC removals were consistent after 20,000 bed volumes. NBDOC and treatment efficiencies experienced a slightly more negative correlation with bed volumes, but after 40,000 bed volumes stayed consistently above 0.20 mg/L and around 20% removal, respectively.

SI Table B.A1: Completed DOC removal data and bed volumes for each sampling event.

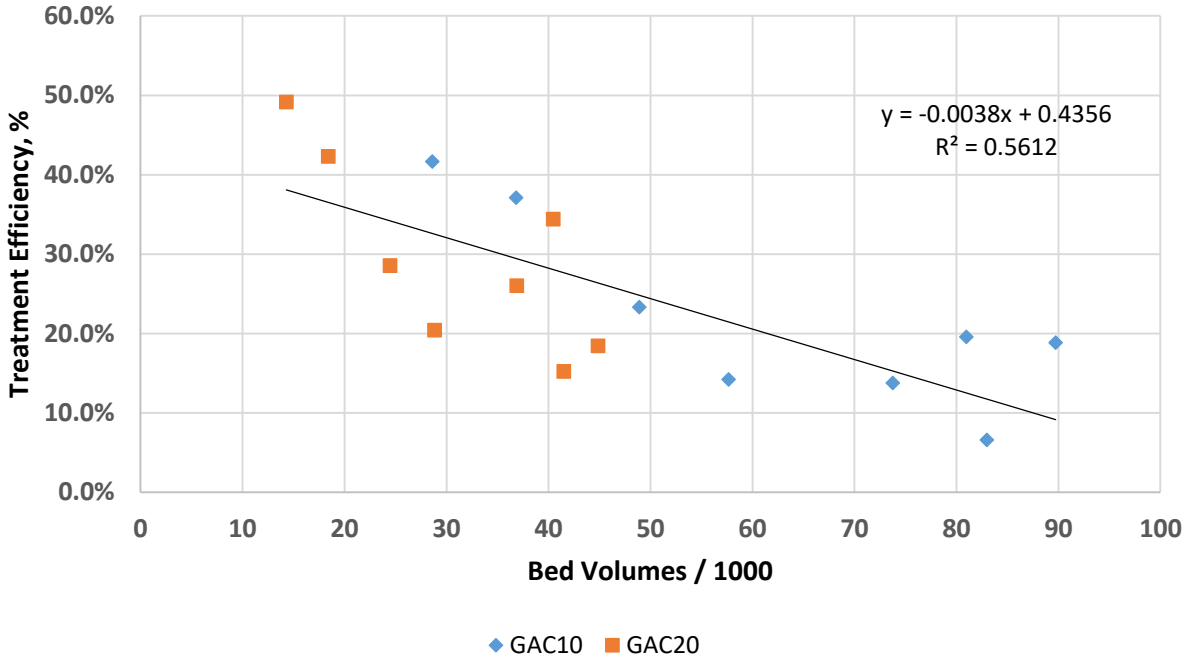
Event	GAC10				GAC20			
	BV	BV/1000	Treatment Removal, mg/L	Treatment Efficiency, %	BV	BV/1000	Treatment Removal, mg/L	Treatment Efficiency, %
S02	28616.2	28.6	3.0	41.7%	14308.1	14.3	2.9	49.2%
S04	36812.8	36.8	2.8	37.1%	18406.4	18.4	2.8	42.3%
S05	48892.0	48.9	0.9	23.3%	24446.0	24.4	1.0	28.6%
S06	57663.8	57.7	0.5	14.2%	28831.9	28.8	0.7	20.4%
S07	73769.4	73.8	0.6	13.8%	36884.7	36.9	1.1	26.0%
S08	80959.4	81.0	0.9	19.6%	40479.7	40.5	1.5	34.4%
S09	82972.6	83.0	0.4	6.6%	41486.3	41.5	0.8	15.2%
S10	89731.2	89.7	1.0	18.9%	44865.6	44.9	0.9	18.4%



SI Figure B.4: Bed volume breakthrough curve with samples from BDOC testing highlighted.

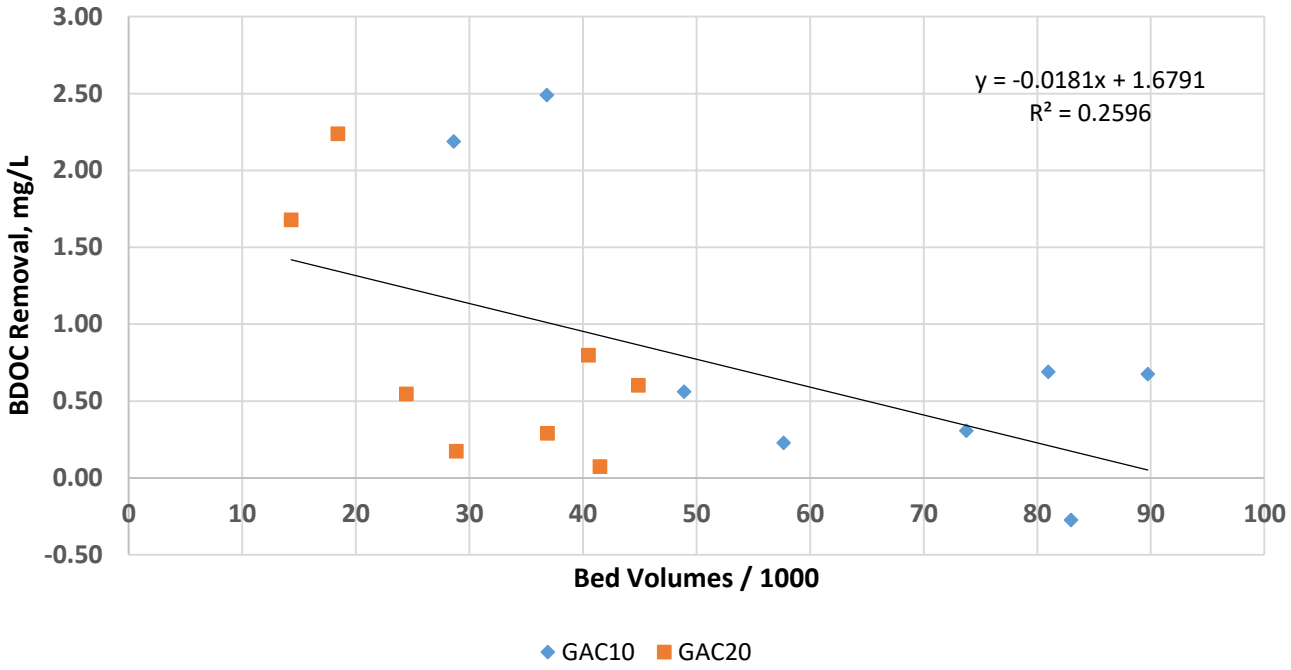


SI Figure B.5: DOC removal achieved during operation benchmarked to bed volumes for both GAC10 and GAC20. Note that the two lowest bed volumes for each GAC contactor were impacted by methanol leaching upstream and therefore may be experiencing elevated values from carry over effects.

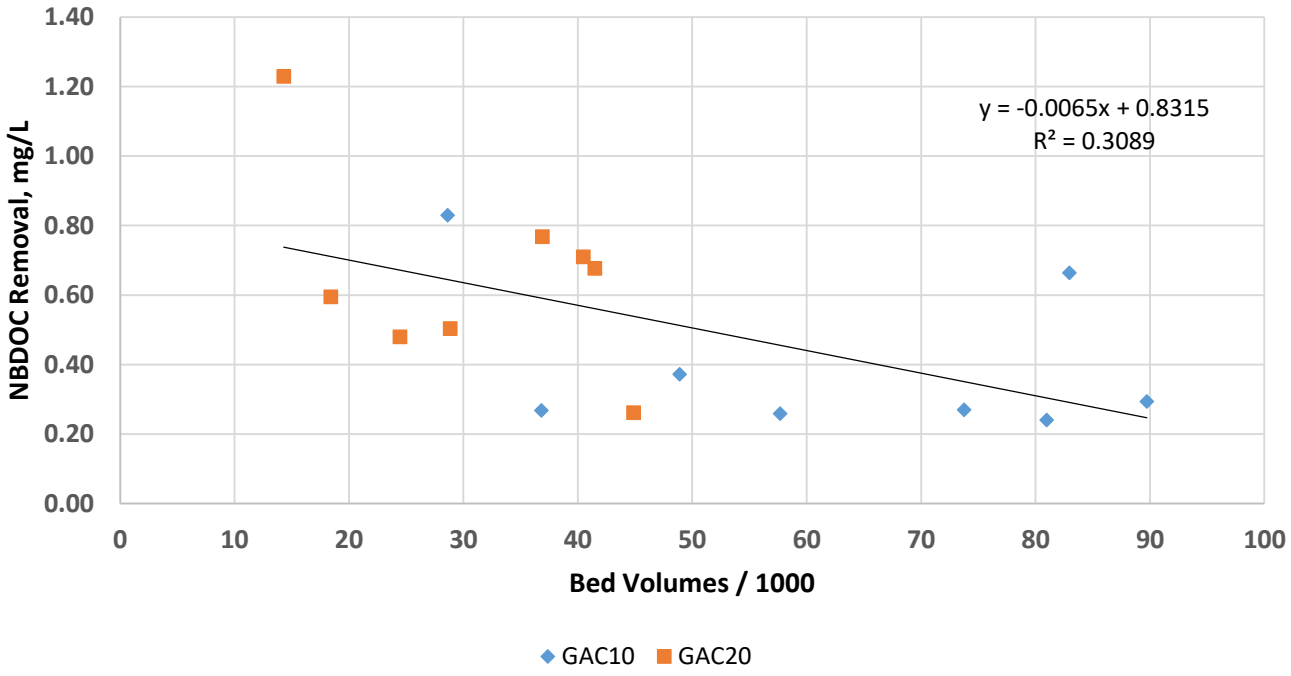


SI Figure B.6: Treatment efficiencies (percentage of DOC removed) achieved during operation benchmarked to bed volumes for both GAC10 and GAC20. Note that the two

lowest bed volumes for each GAC contactor were impacted by methanol leaching upstream and therefore may be experiencing elevated values from carry over effects.

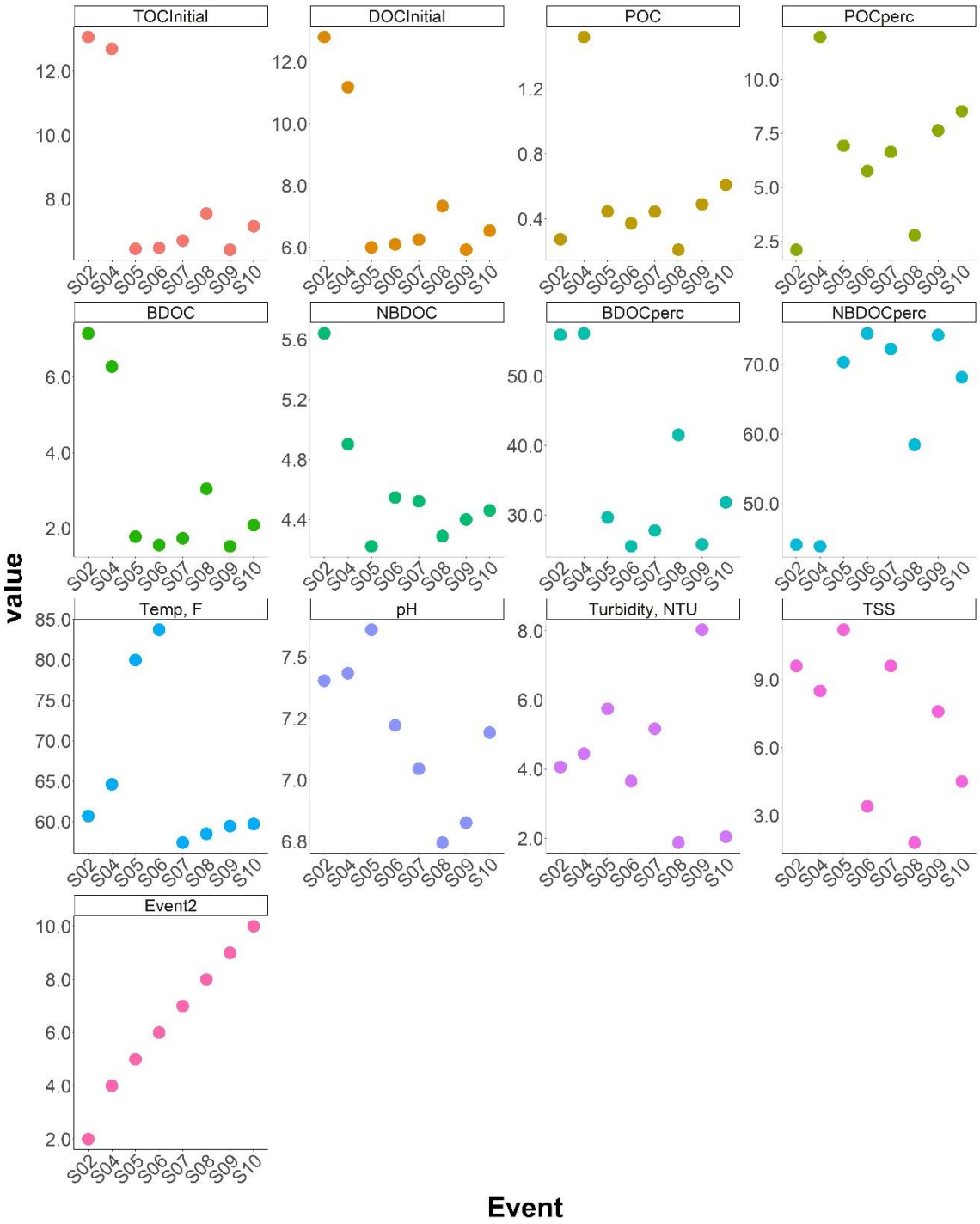


SI Figure B.7: BDOC (biodegradable dissolved organic carbon) removal achieved during operation benchmarked to bed volumes for both GAC10 and GAC20. Note that the two lowest bed volumes for each GAC contactor were impacted by methanol leaching upstream and therefore may be experiencing elevated values from carry over effects.

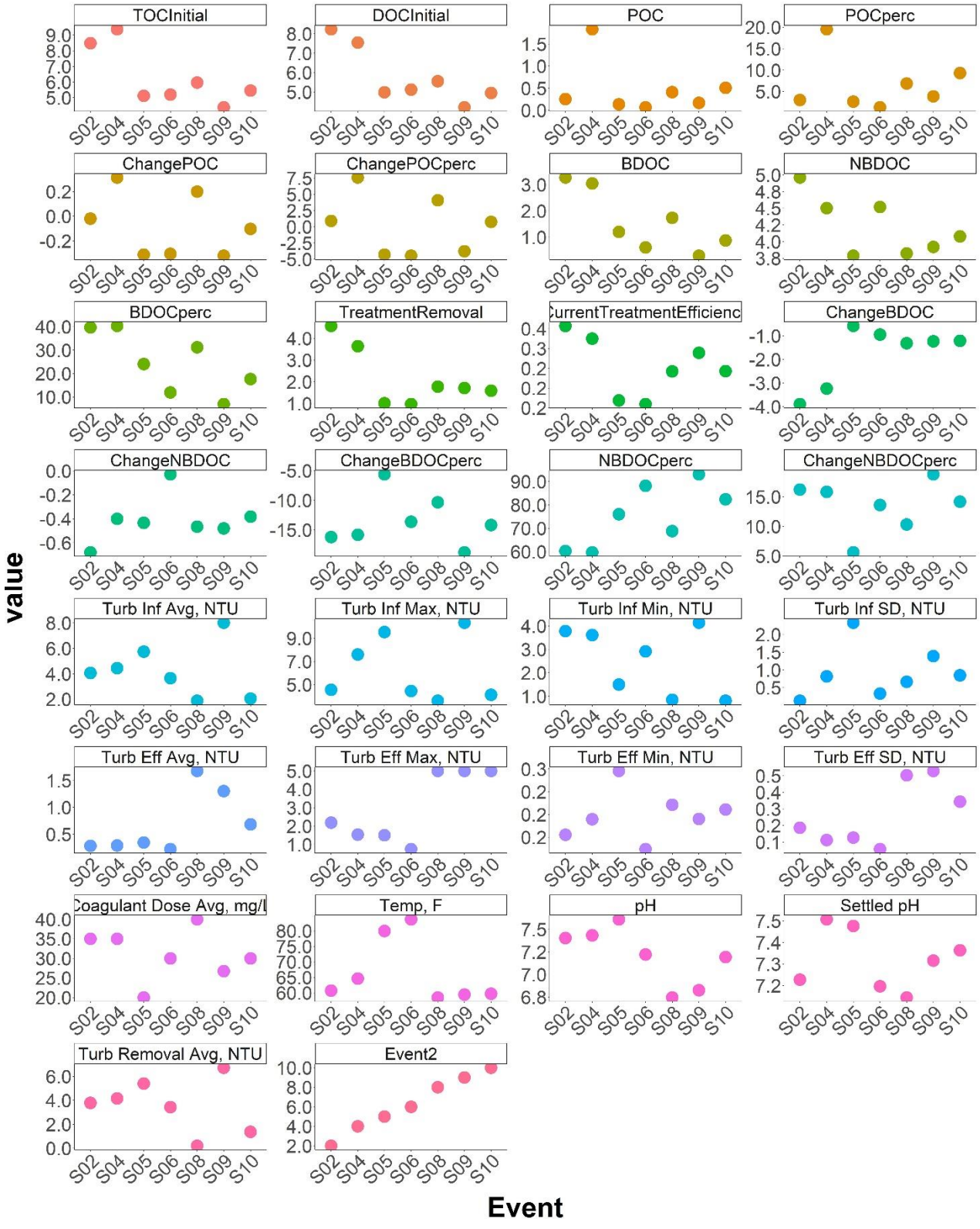


SI Figure B.8: NBDOC (non-biodegradable dissolved organic carbon) removal achieved during operation benchmarked to bed volumes for both GAC10 and GAC20. Note that the two lowest bed volumes for each GAC contactor were impacted by methanol leaching upstream and therefore may be experiencing elevated values from carry over effects.

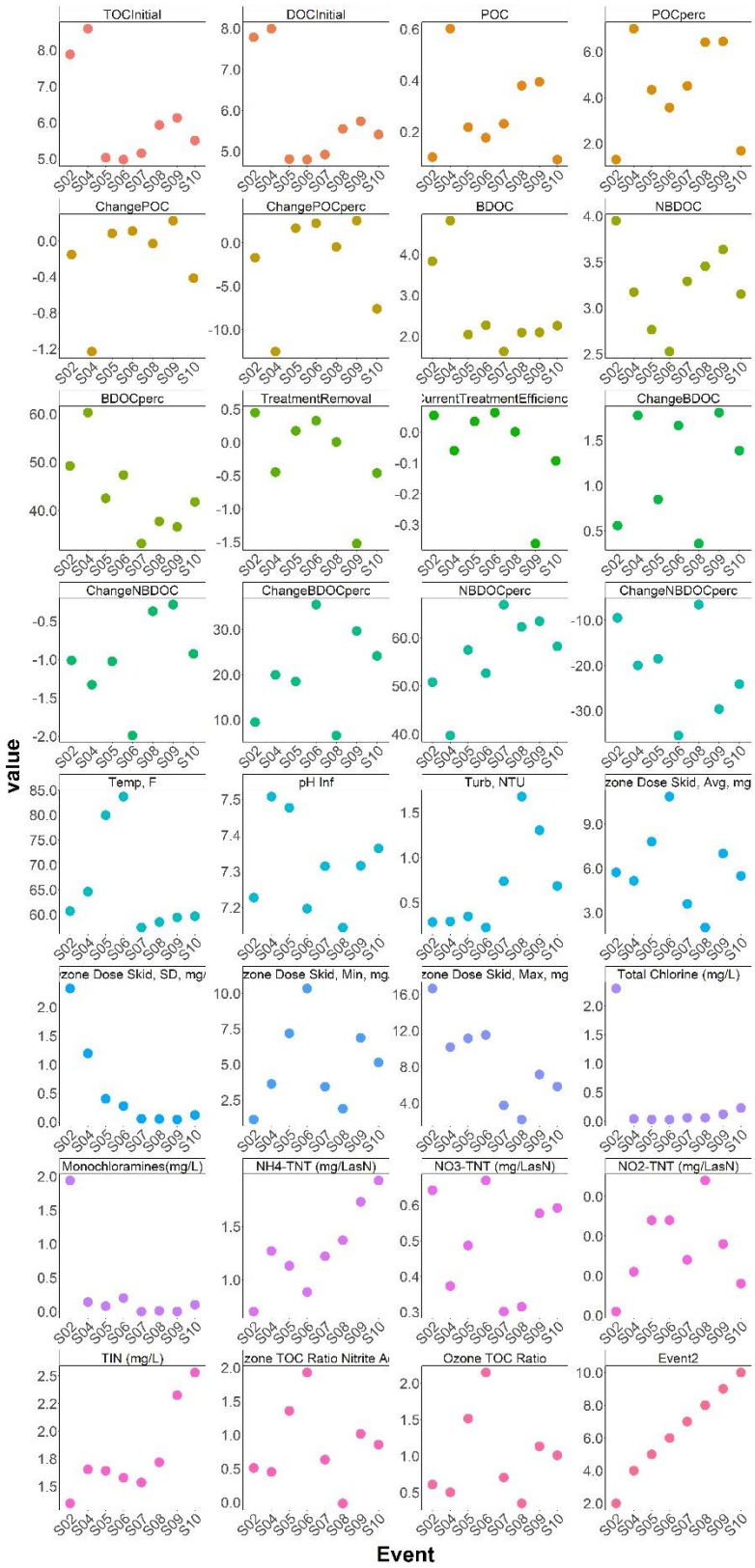
Subsection 2: Operational Conditions and Water Quality Data



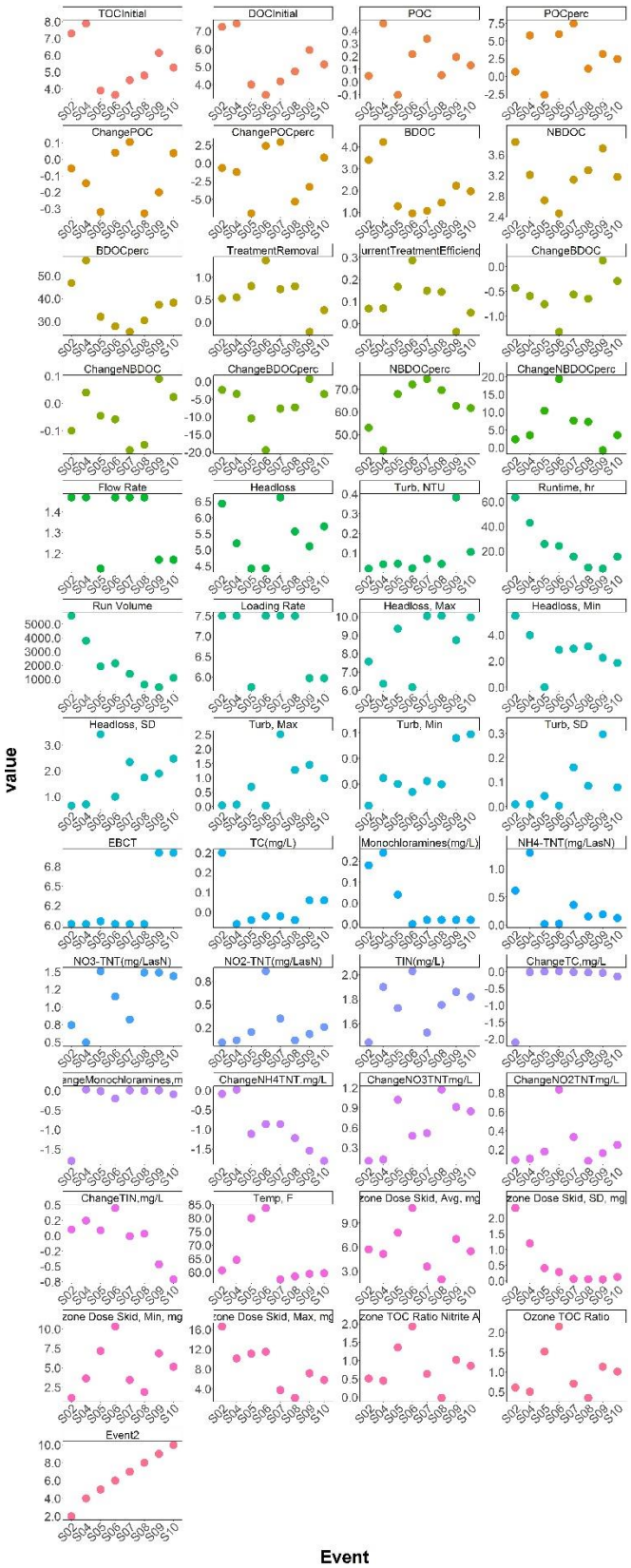
SI Figure B.9: Pilot influent supplemental data.



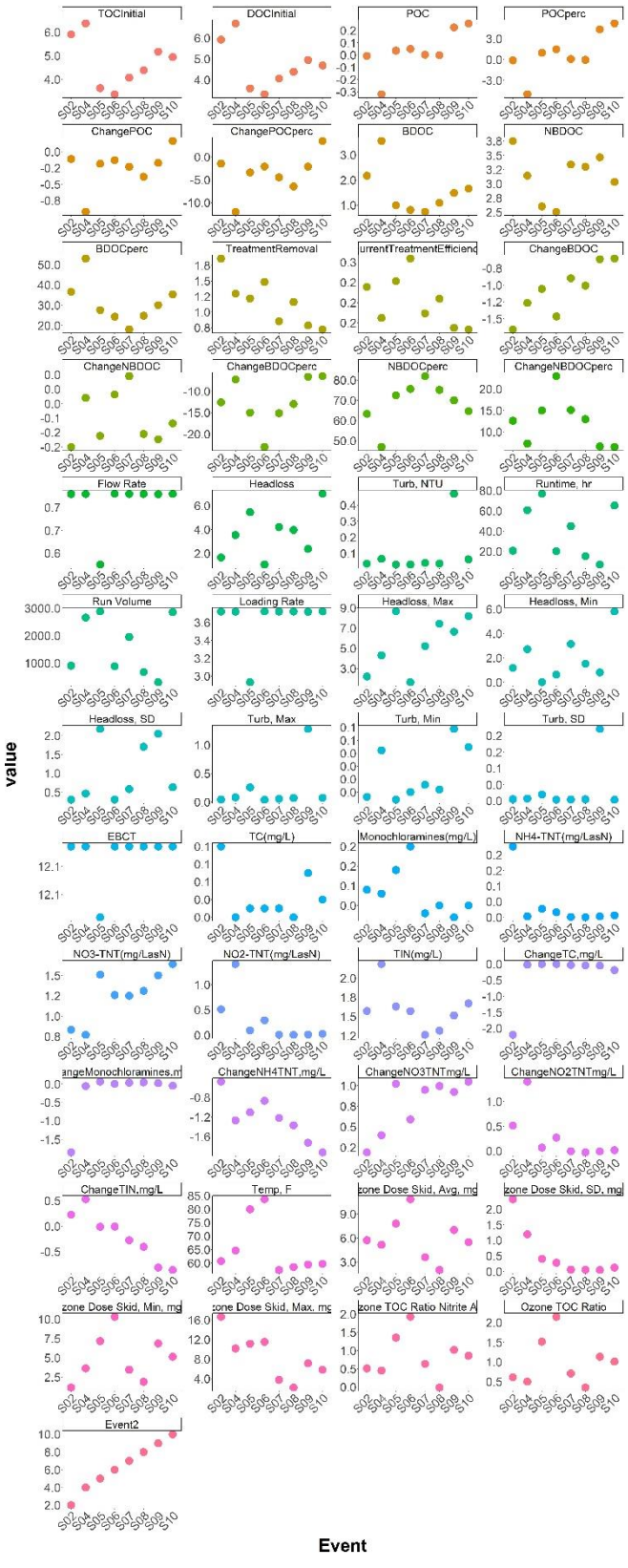
SI Figure B.10: Floc-sed supplemental data.



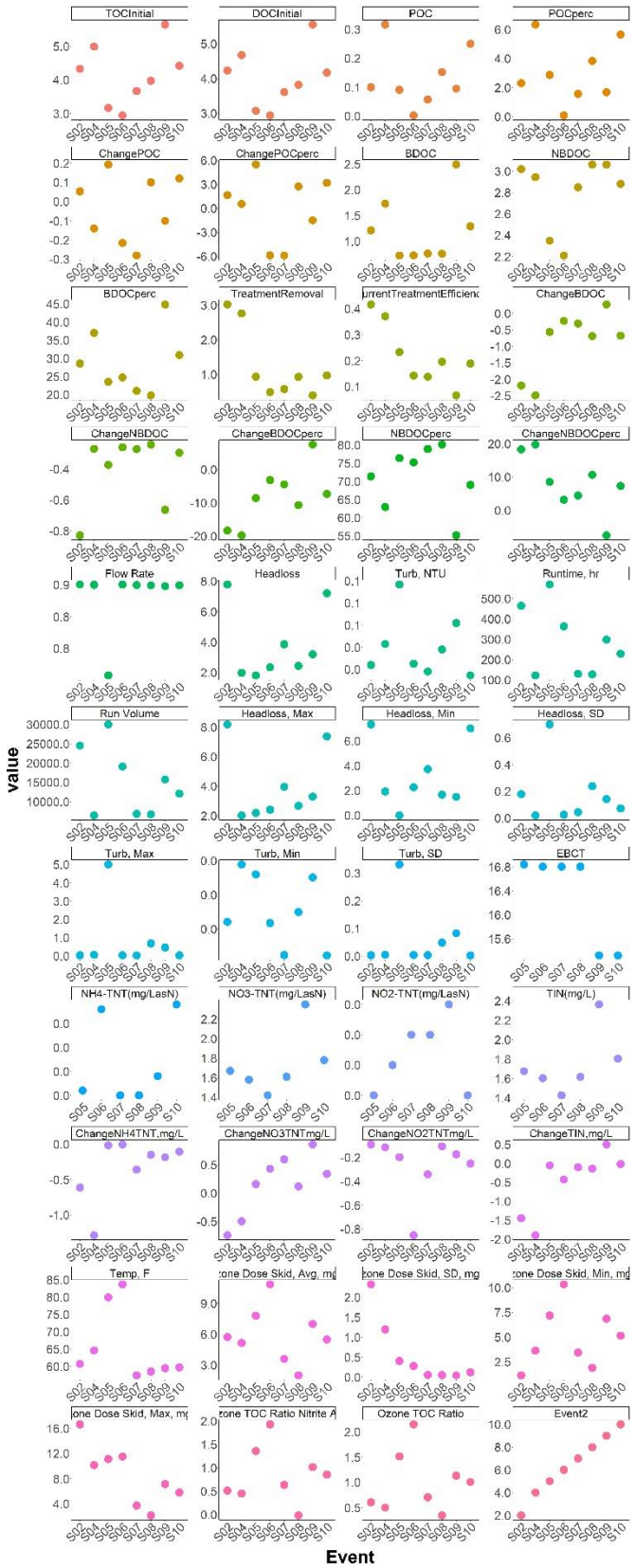
SI Figure B.11: Ozone supplemental data.



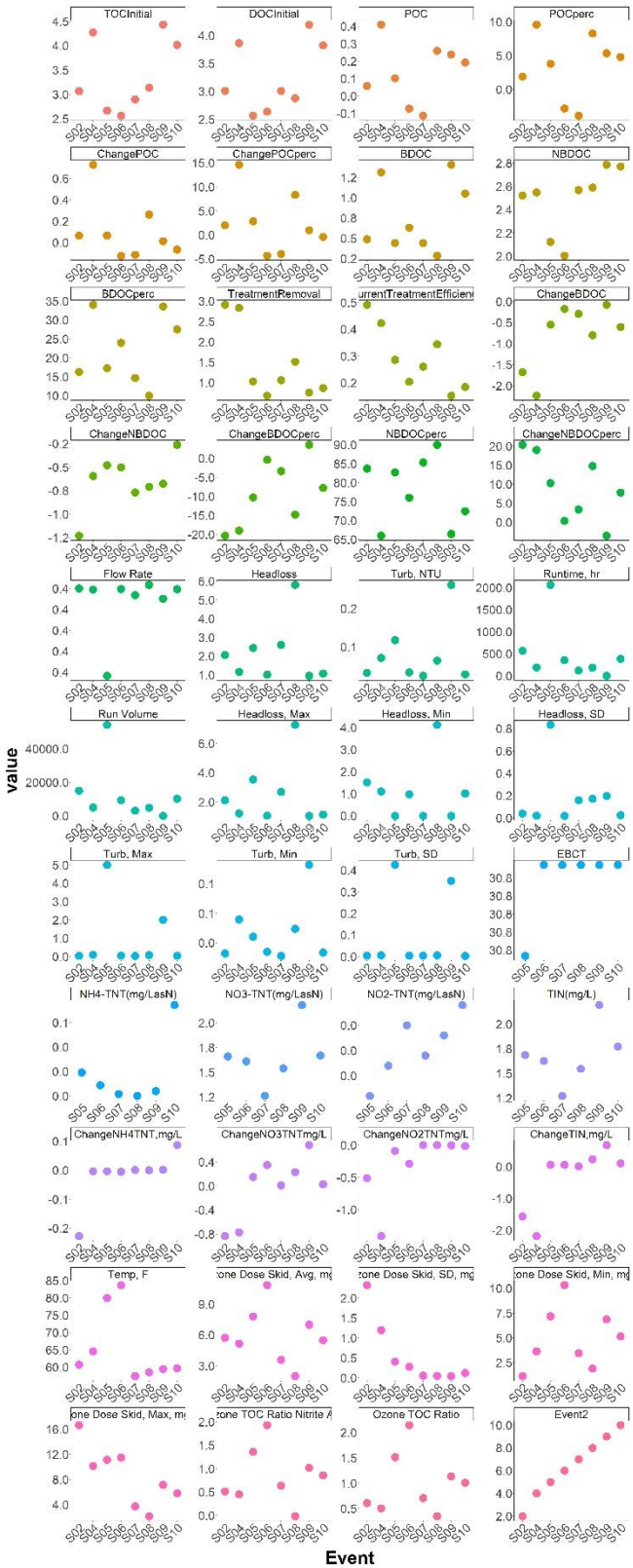
SI Figure B.12: BAC-5 supplemental data.



SI Figure B.13: BAC-10 supplemental data.



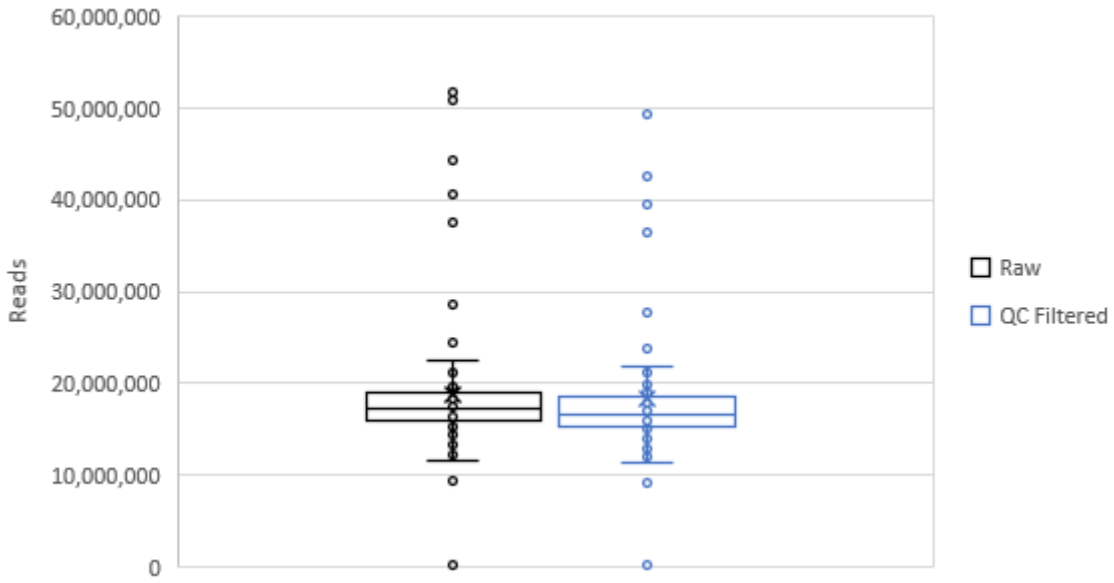
SI Figure B.14: GAC-10 supplemental data.



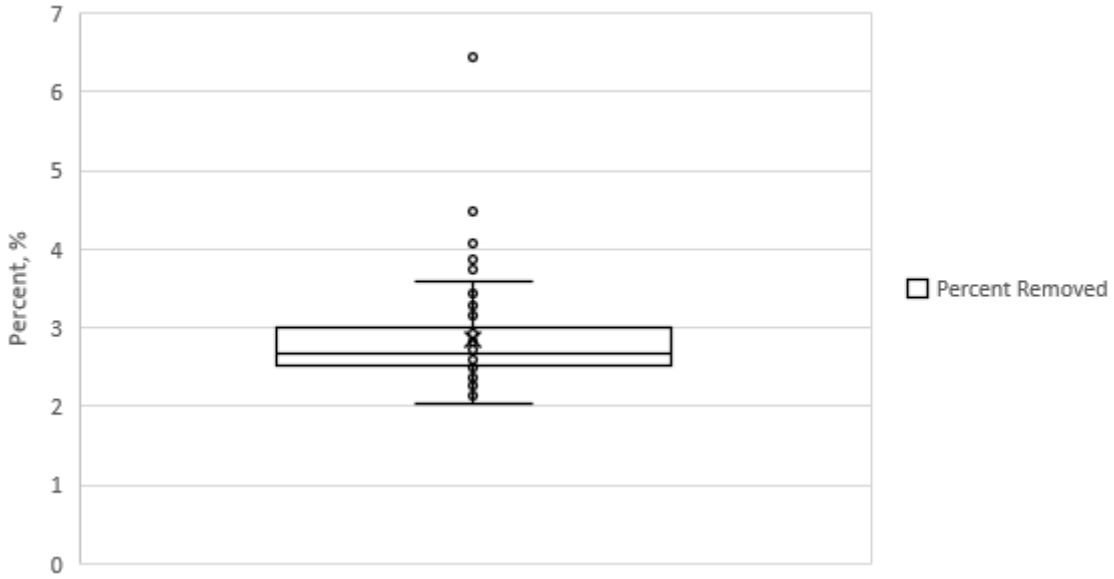
SI Figure B.15: GAC-20 supplemental data.

APPENDIX C: BACKGROUND DATA ON SHOTGUN METAGENOMIC SEQUENCING

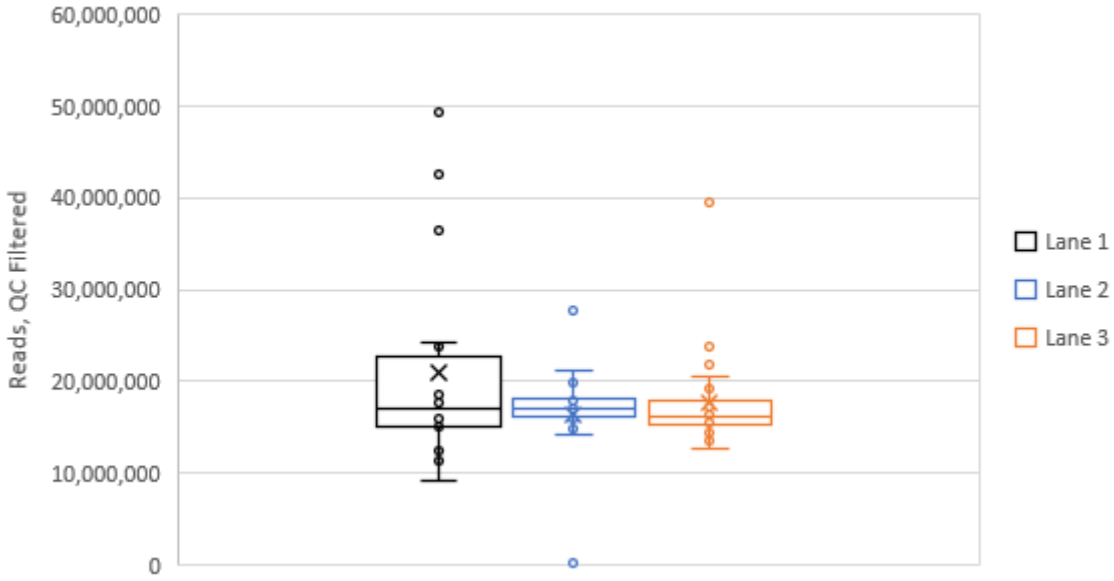
All 72 samples had their DNA concentrations quantified using Qubit2.0 and were then broken up into three lanes where samples were diluted to match their DNA concentrations. Using this tiered approach, we avoided having to dilute our highest concentrated samples down to our lowest concentrated samples. SI Figure C.1 shows a box and whisker plot of raw reads and QC filtered reads for all 72 samples while SI Figure C.2 shows the percentage of reads removed by QC filtering. SI Figure C.3 shows QC filtered read variability between lanes while SI Figure C.4 shows QC filtered read variability between sampling locations independent of the lane they were sequenced on. These graphs indicate that read quality was good (consistent, low percentage of reads removed during QC filtering) and comparable between lanes and samples (median ~ 17,000,000 for all). Outliers were also seemingly random and not a function of sample location or lane.



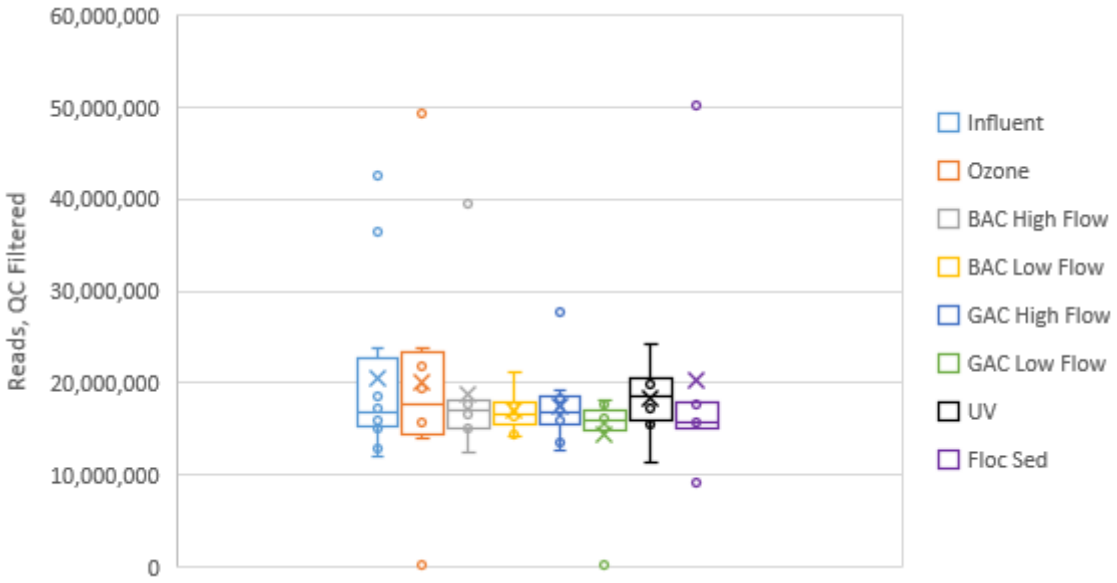
SI Figure C.1: Number of raw and QC filtered sequencing reads for all shotgun metagenomic samples.



SI Figure C.2: Percentage of reads removed during the QC filtering process for each sample.



SI Figure C.3: Number of QC filtered sequencing reads for all shotgun metagenomic samples grouped by lane.



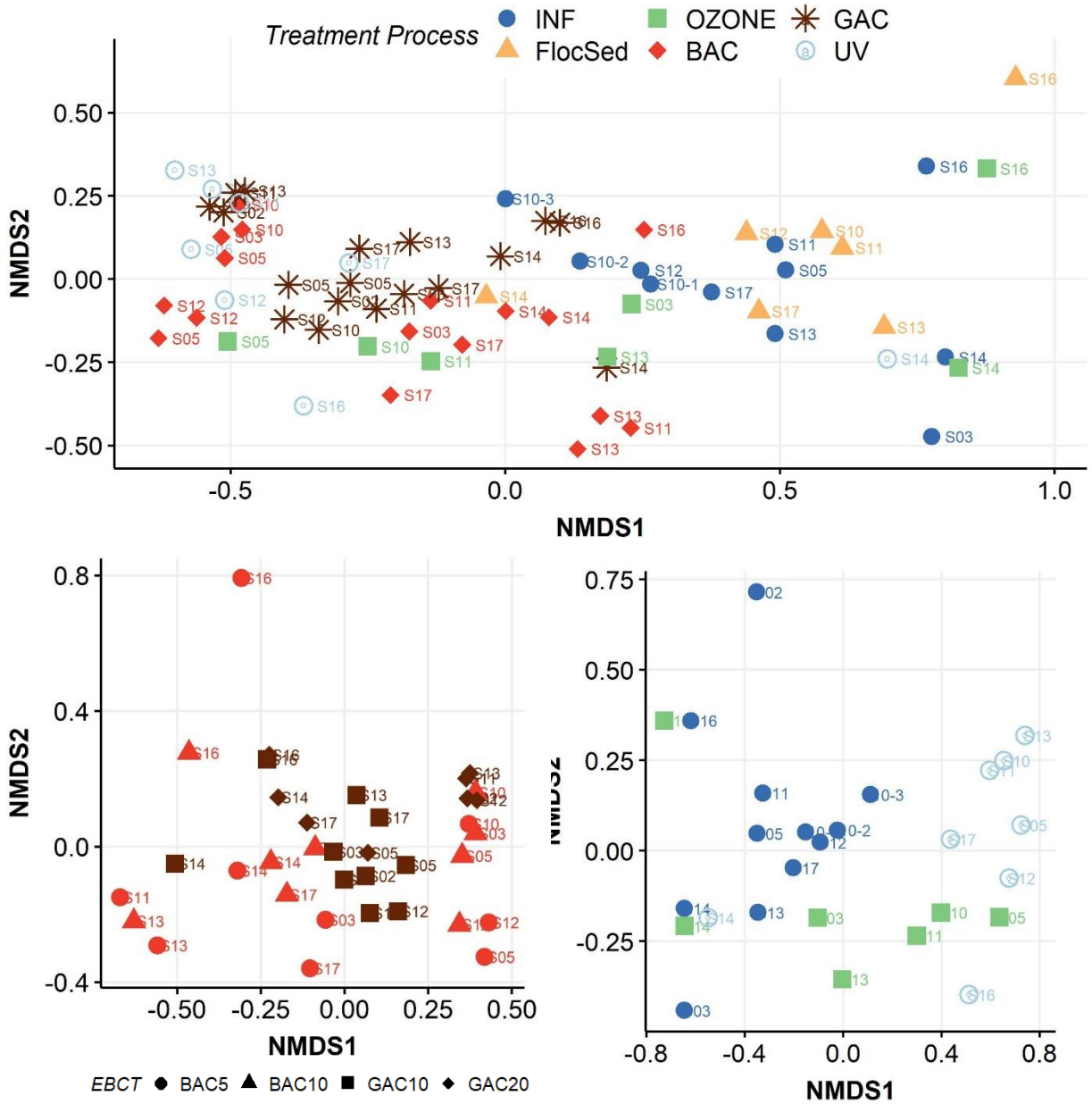
SI Figure C.4: Number of QC filtered sequencing reads for all shotgun metagenomic samples grouped by location.

APPENDIX D: SUPPLEMENTAL INFORMATION FOR CHAPTER 3

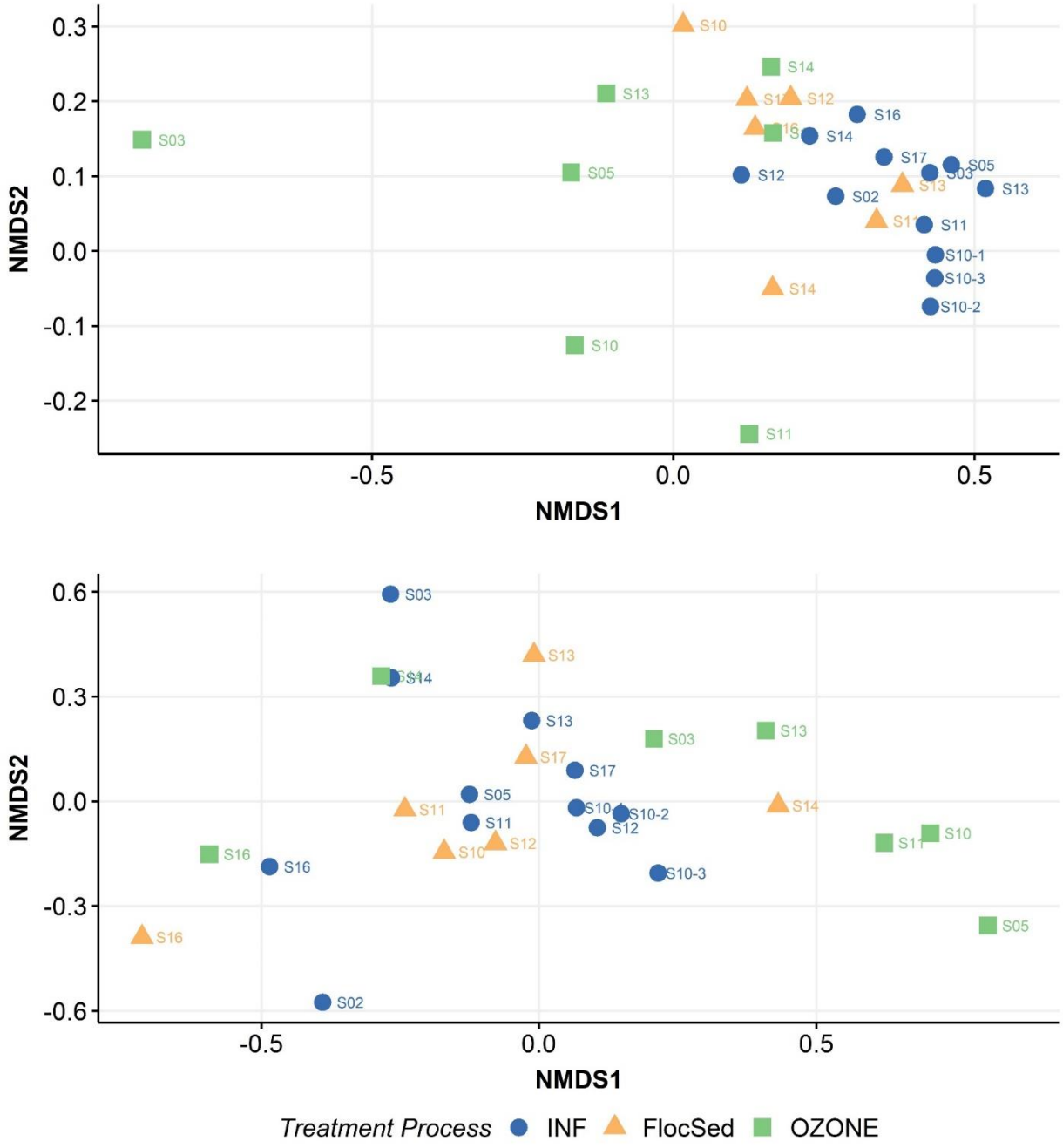
SI Table D.1: Overview of Collected Samples, Operational Conditions, and Upset Events

Sampling Event	Date Sampled or Started	Specific Operational Conditions and Changes	System Upsets and Anomalies	Submitted For Metagenomic Sequencing
-	Before 09/01/2016	**Chlorinated Backwashed Start**	-	
-	09/01/2016	**Monochloramine Dosing Start**	-	
SS-Startup	09/03/2016	-	Limited Replication	
SS-01	11/12/2016	-	-	
SS-02	12/14/2016	-	-	
SS-03	01/19/2017	-	-	
SS-04	02/17/2017	-	-	
-	02/22/2017	**Chlorine Quenching Started (Bisulfite)**	-	
-	03/20/2017	**Chlorinated Backwash Stopped**	-	
SS-05	3/22/2017	-	-	
-	04/01/2017	-	Media Loss BAC5, supplemented with media from BAC10, flow rate adjusted to maintain EBCT	
SUMediaLoss-06	04/12/2017	-	See above	
SS-07	05/17/2017	-	-	
SS-08	06/14/2017	-	-	
SS-09	07/20/2017	-	-	
SS-10	08/23/2017	-	-	
-	09/18/2017	**Nutrient Additions Start** Ammonia (NH4)2SO4 0.5 mg/L as N, Phosphorous H3PO4 0.1 mg/L as P	-	
SSNutrients-11	10/10/2017	1:1 O3:TOC, Thiosulfate Quench, Additional Nutrients (N, P)	-	
-	10/10/2017	**Chiller Start (15° C)**	-	
SSChiller-12	11/14/2017	1:1 O3:TOC, Chilled (15° C), Thiosulfate	-	

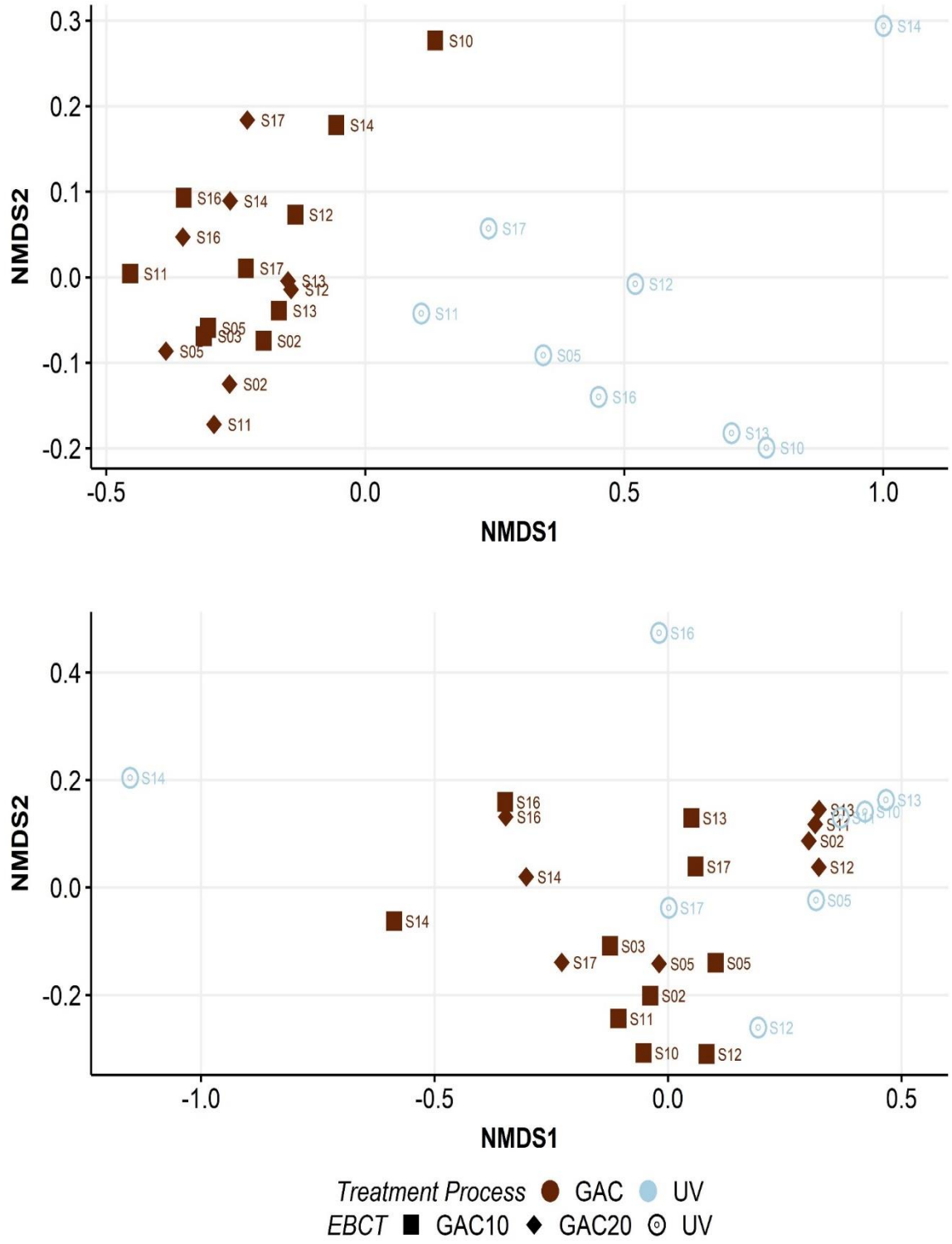
		Quench, Additional Nutrients (N, P)	
SSMediumOzone-13	12/12/2017	0.5:1 O3:TOC, Chilled (15° C), Thiosulfate Quench, Additional Nutrients (N, P)	-
SSLowOzone-14	01/31/2018	0.25:1 O3:TOC, Chilled (15° C), Thiosulfate Quench, Additional Nutrients (N, P)	-
-	02/02/2018	**Return to Bisulfite Quenching**	-
-	02/02/2018	**Filter Aid Polymer Added**	-
SUHighTurbidity-15	02/14/2018	1:1 O3:TOC, Chilled (15° C), Bisulfite Quench, Additional Nutrients (N, P), Filter Polymer Added	Low coagulant dosing (stock issues) lead to large issues with Floc Sed performance. High turbidity event.
-	03/6/2018	**Chlorinated Backwashes Start**	-
SSHighTurbidity-16	03/19/2018	0.8:1 O3:TOC, Chilled (15° C), Bisulfite Quench, Additional Nutrients (N, P), Filter Polymer Added, Chlorine Residual in Backwash	Pilot seeing higher than normal turbidity with elevated nitrite (Ozone consumption)
-	-	**Residual Chlorine Loading Start**	-
SSNoQuench-17	05/19/2018	0.8:1 O3:TOC, Chilled (15° C), No Quench, Additional Nutrients (N, P), Filter Polymer Added	-



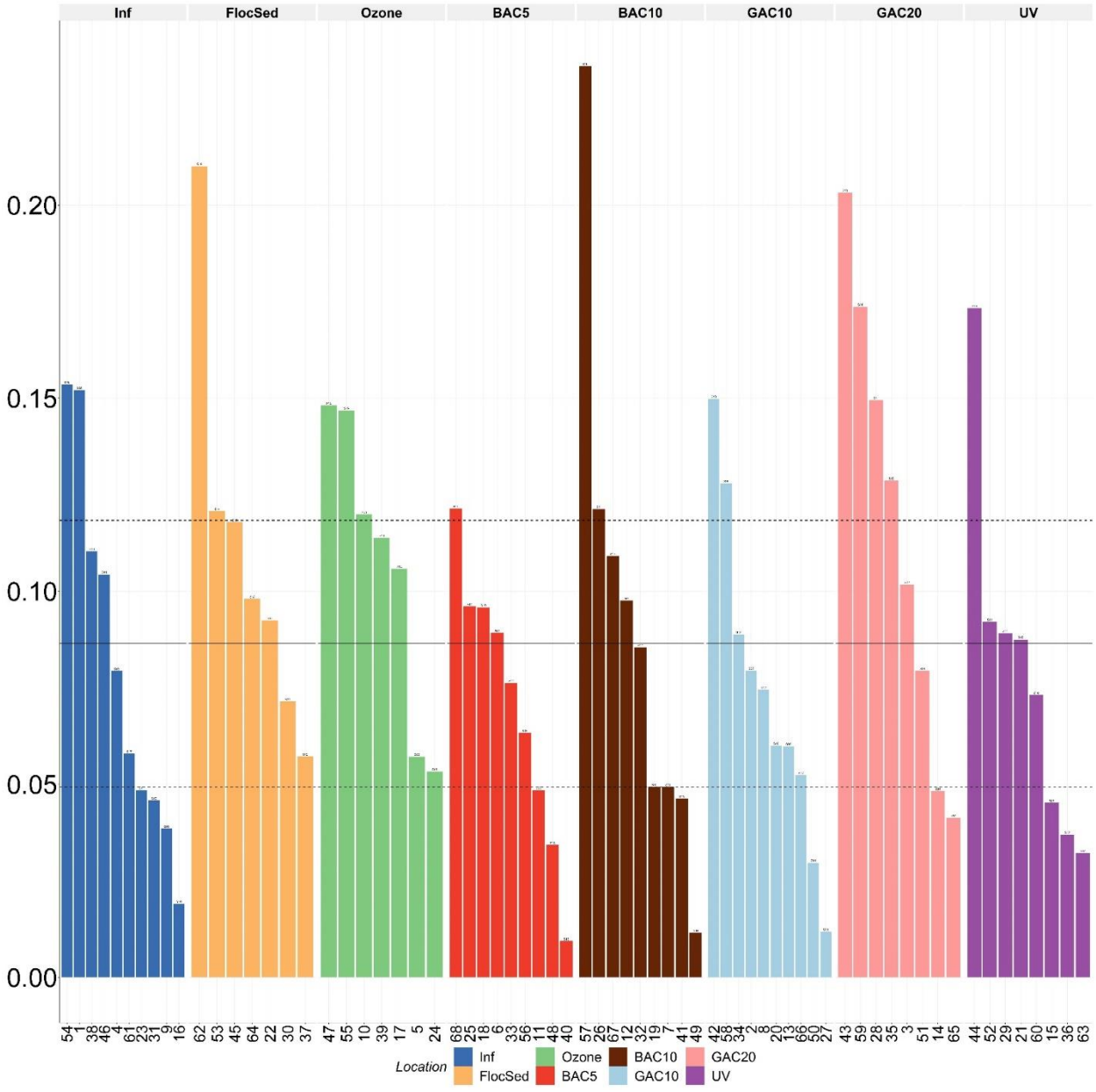
SI Figure D.1: NMDS analysis of resistome (based on relative abundances) for (A) all treatment processes (ANOSIM, p-value<0.05, r-stat = 0.42), (B) BAC filtration and GAC contracting at tested EBCTs (ANOSIM, p-value<0.05, r-stat = 0.14), and (C) disinfection processes (ozone effluent and UV effluent) referenced to treatment train influent (ANOSIM, p-value<0.05, r-stat = 0.40).



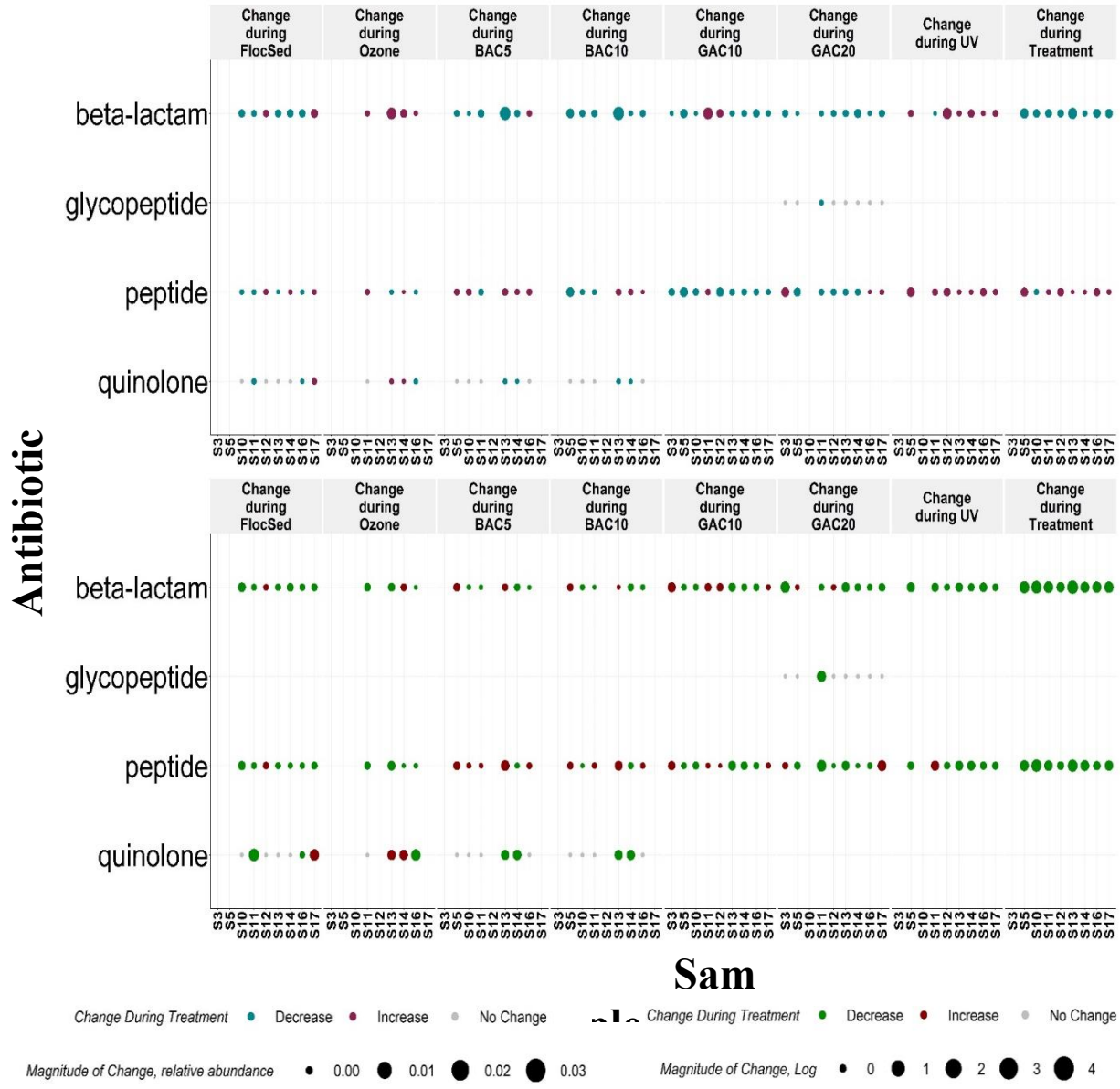
SI Figure D.2: NMDS analysis of resistome for the influent, flocsed, and ozone samples for (A) calculated absolute abundances and (B) relative abundances.



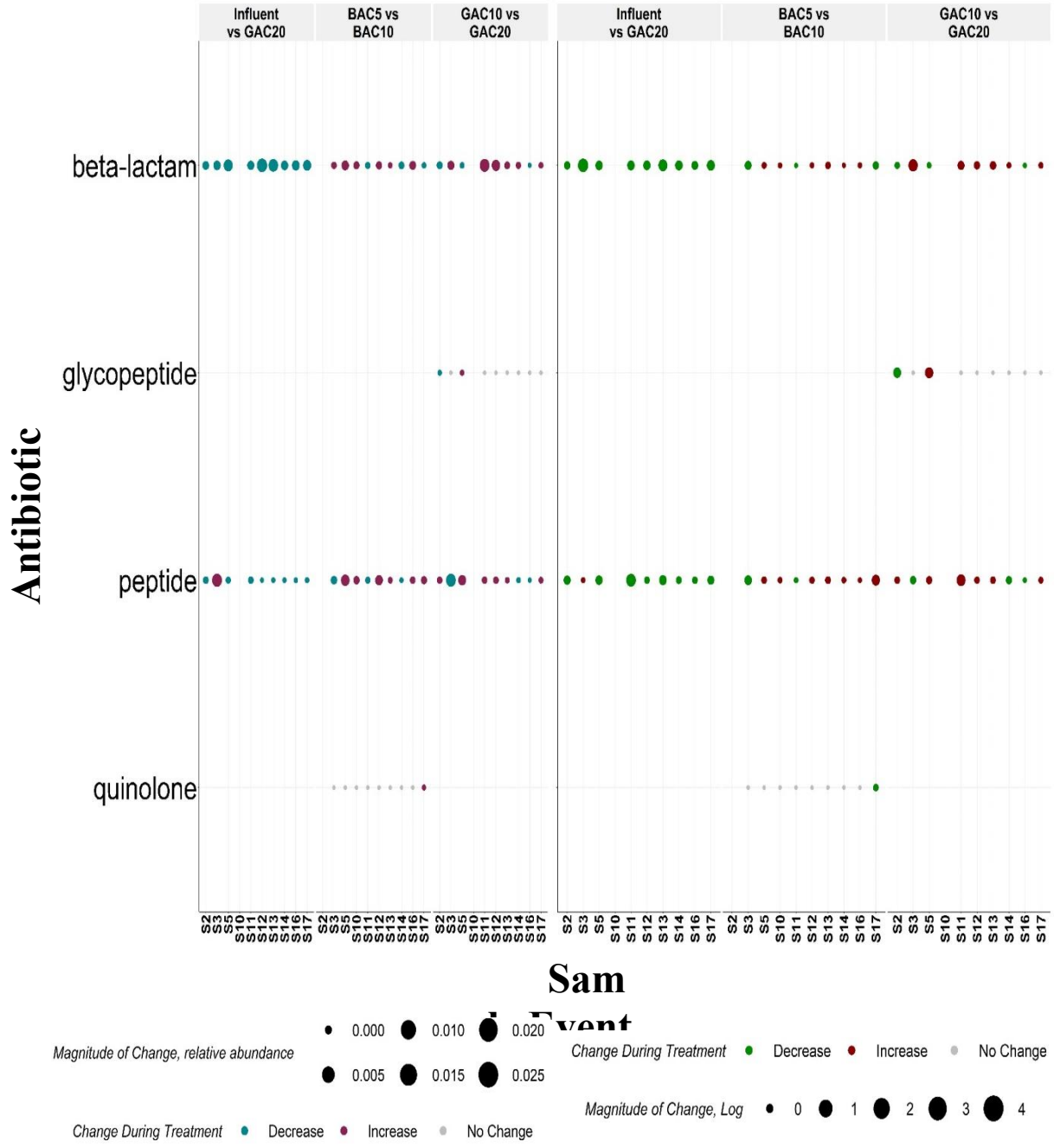
SI Figure D.3: NMDS analysis of resistome for GAC effluent and UV effluent for (A) calculated absolute abundances and (B) relative abundances.



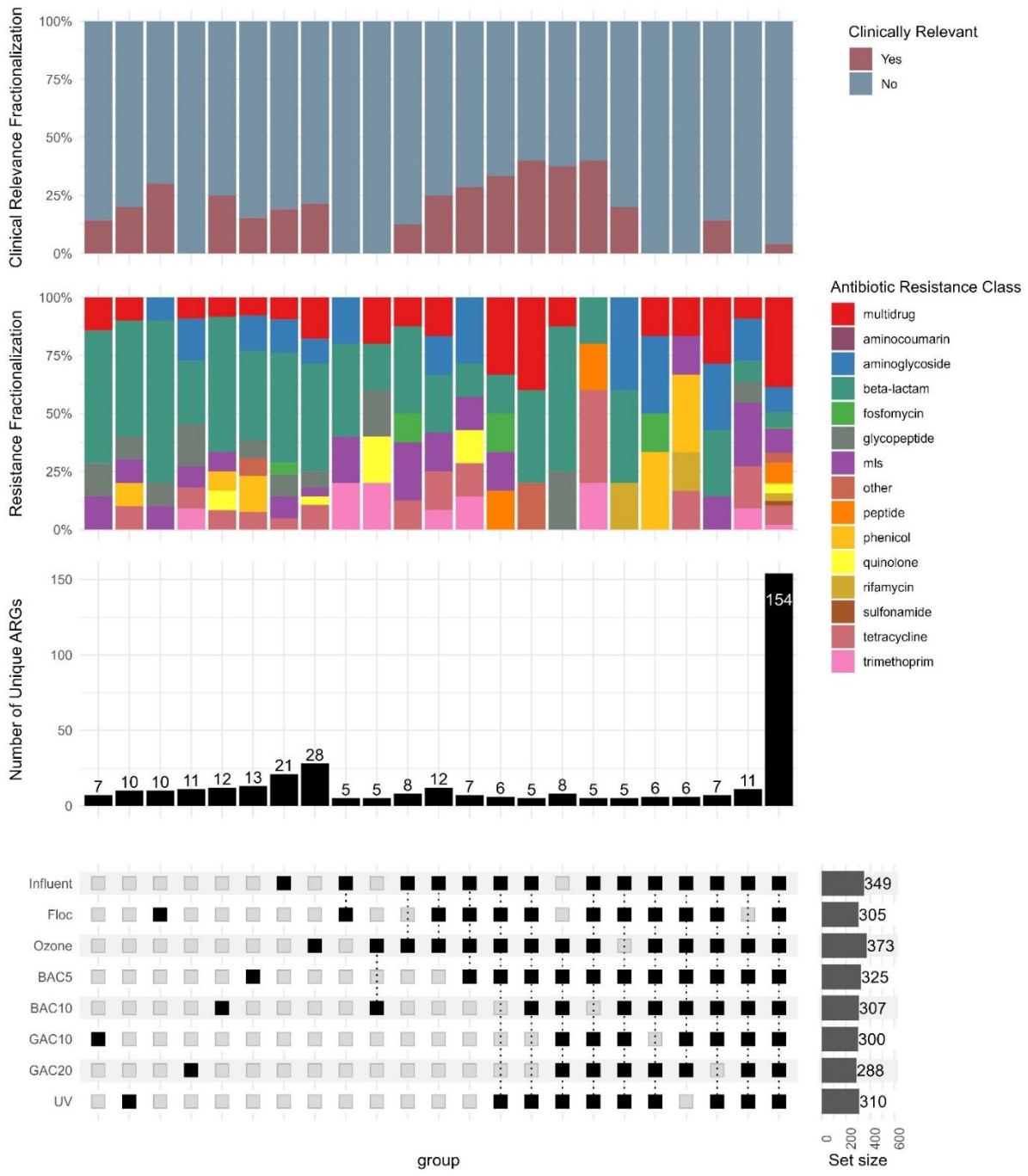
SI Figure D.4: Procrustes analyses were performed on the resistome and taxonomic NMDSs of all samples from each stage of treatment. Residuals associated with the Procrustes analysis are graphically presented above.



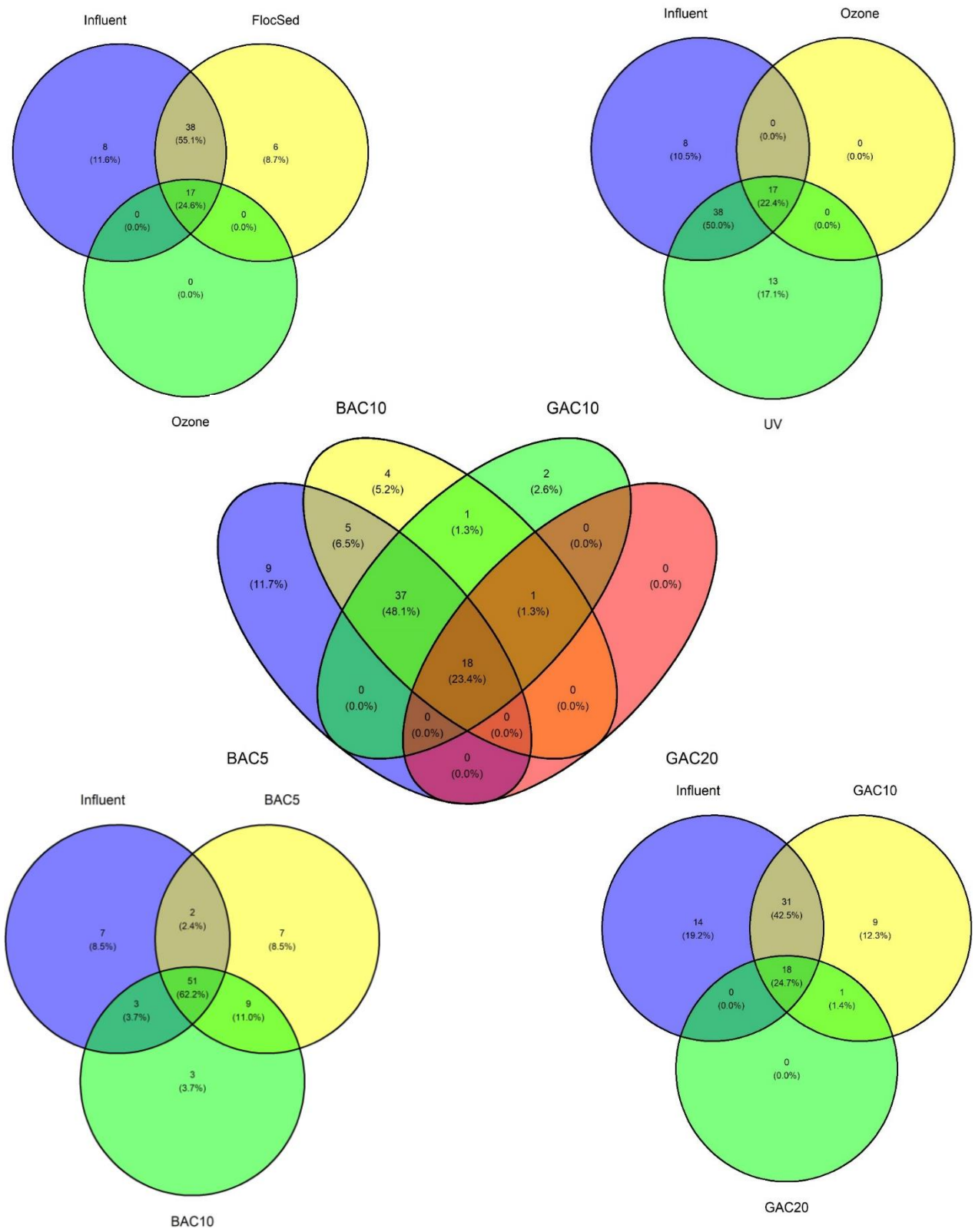
SI Figure D.5: Magnitude of change for each antibiotic resistance class at each sample location and throughout the entire treatment train for (A) difference in 16S rRNA normalized abundances and (B) log change of the calculated absolute abundances. Only clinically relevant genes.



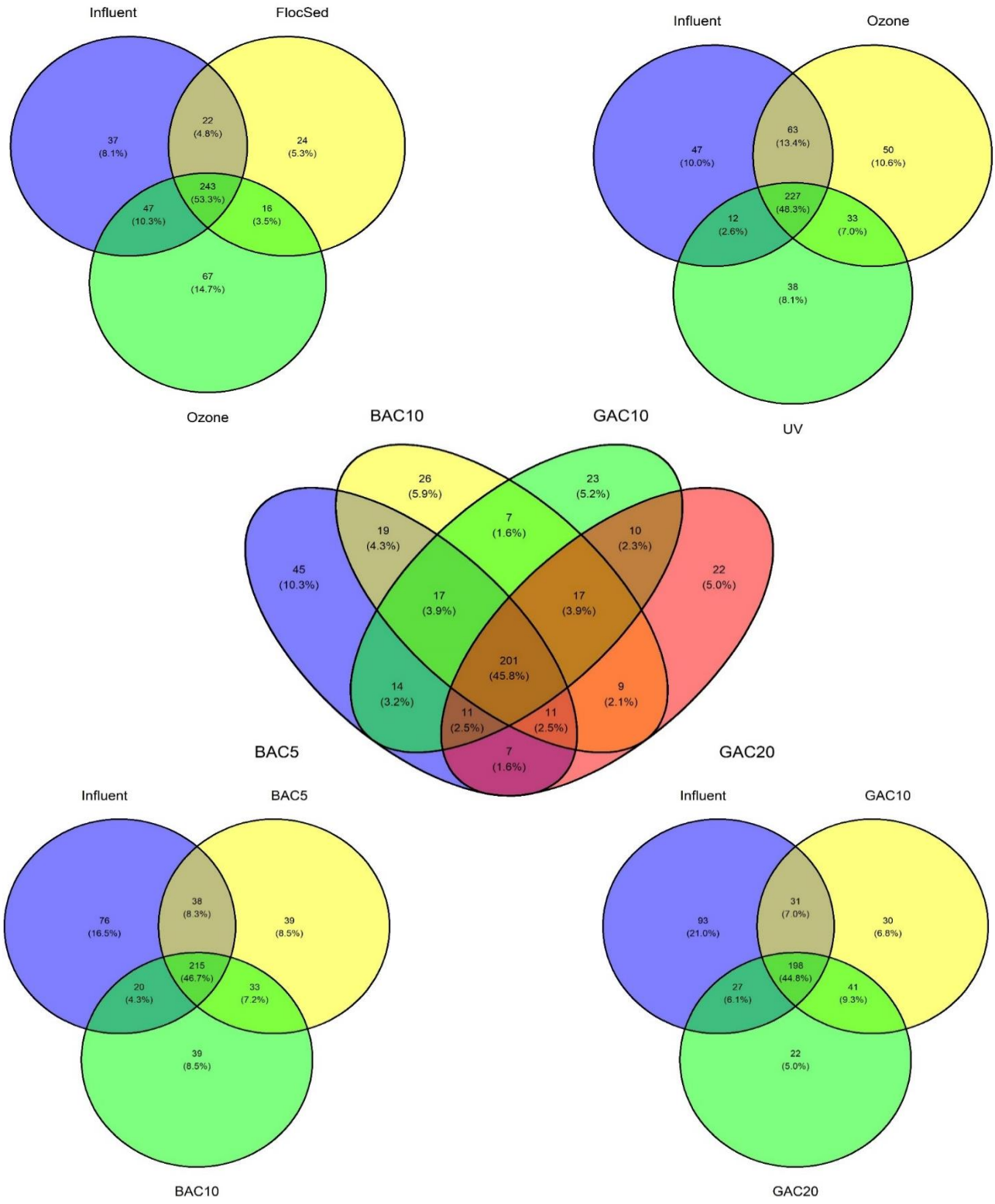
SI Figure D.6: Magnitude of change for each antibiotic resistance class between BACs and GACs operated at different EBCTs and BAF for (A) difference in 16S rRNA normalized abundances and (B) log change of the calculated absolute abundances. Only clinically relevant genes.



SI Figure D.7: Provides a breakdown of all unique ARGs identified throughout treatment and the relationships those unique ARGs have with multiple treatment processes. ARGs are also classified by their clinical relevance and resistance classification. Set size provides the total number of unique ARGs for the associated treatment process.

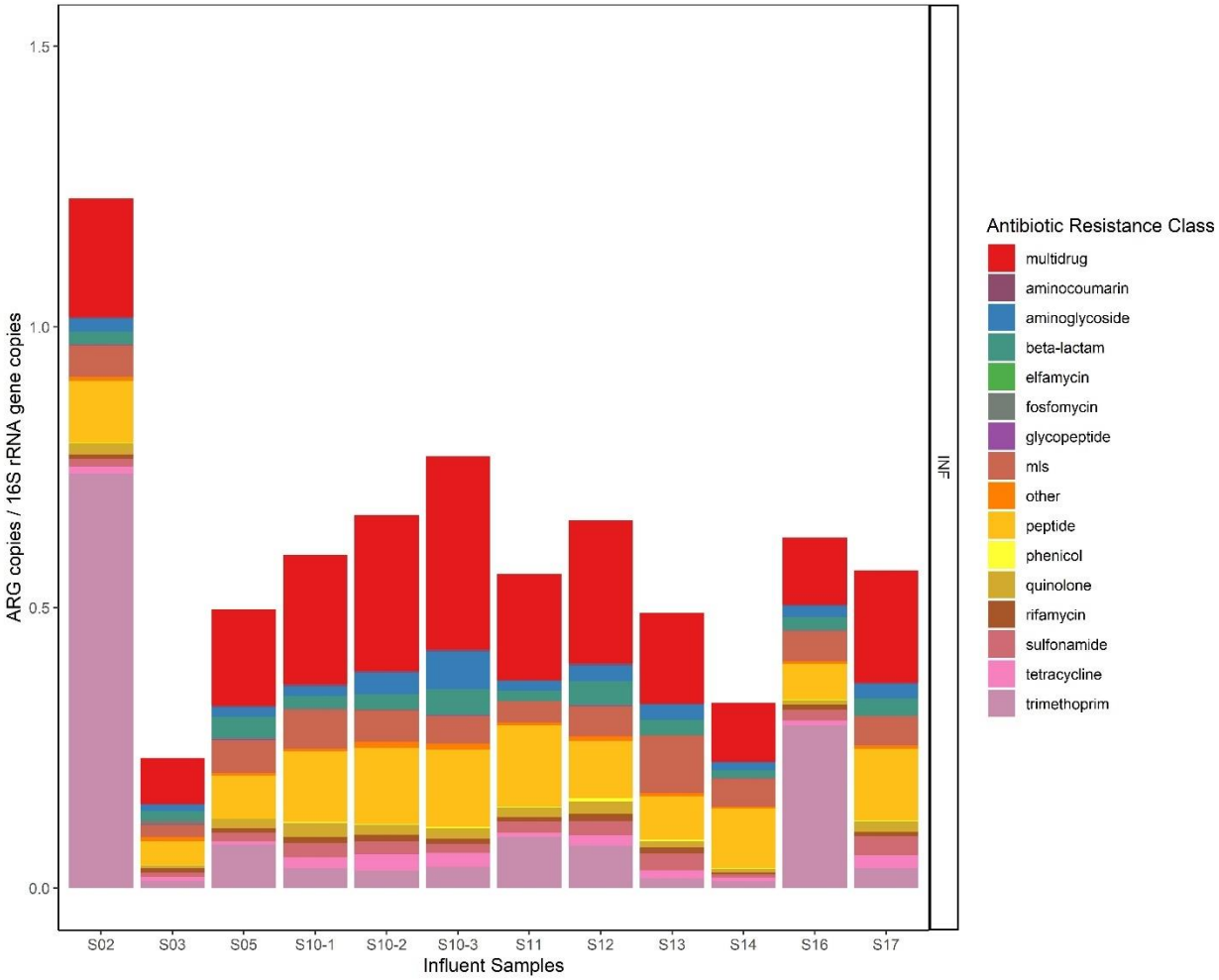


SI Figure D.8: Relationships of core ARGs between specific comparisons of interest.

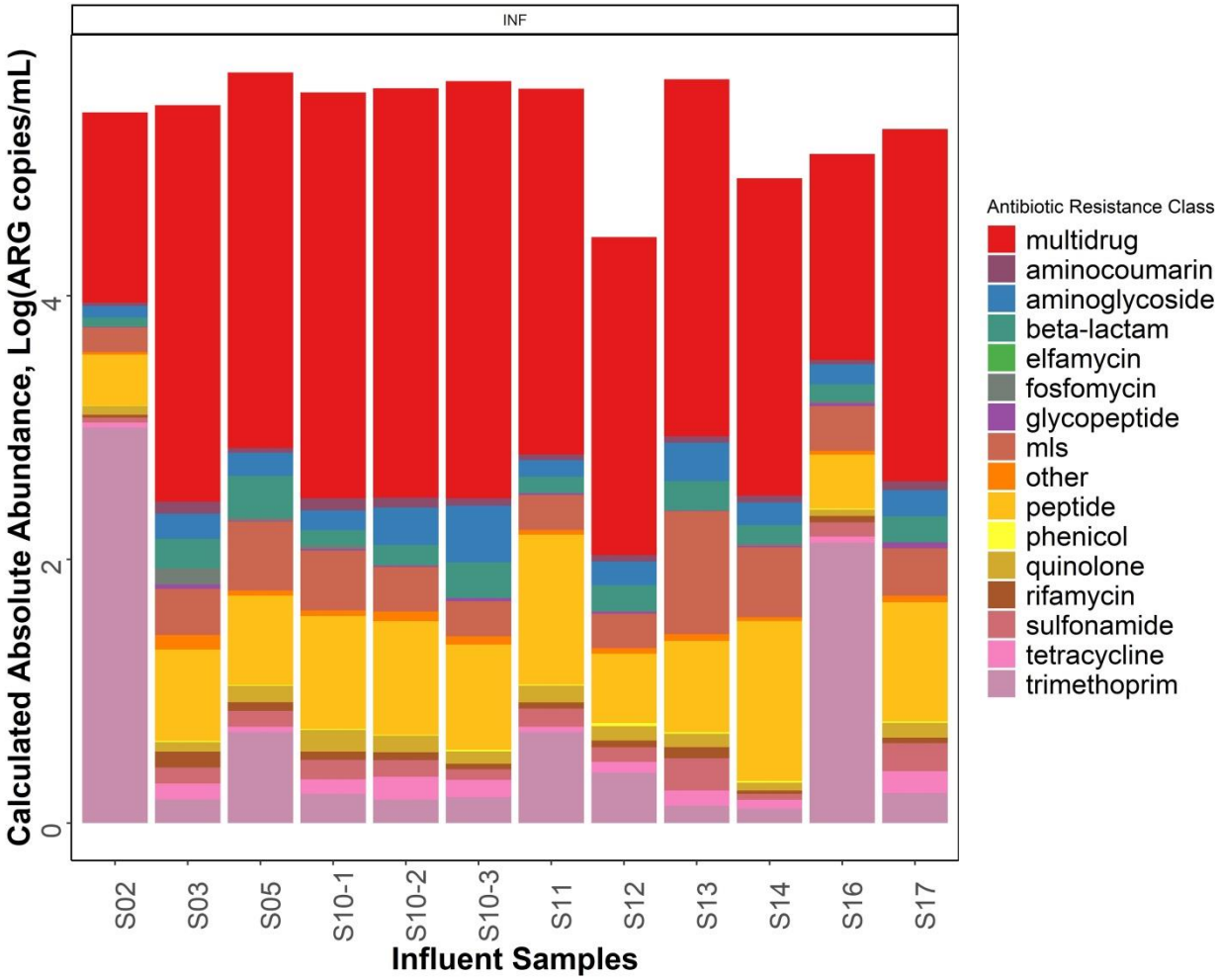


SI Figure D.9: Relationships of unique ARGs between specific comparisons of interest.

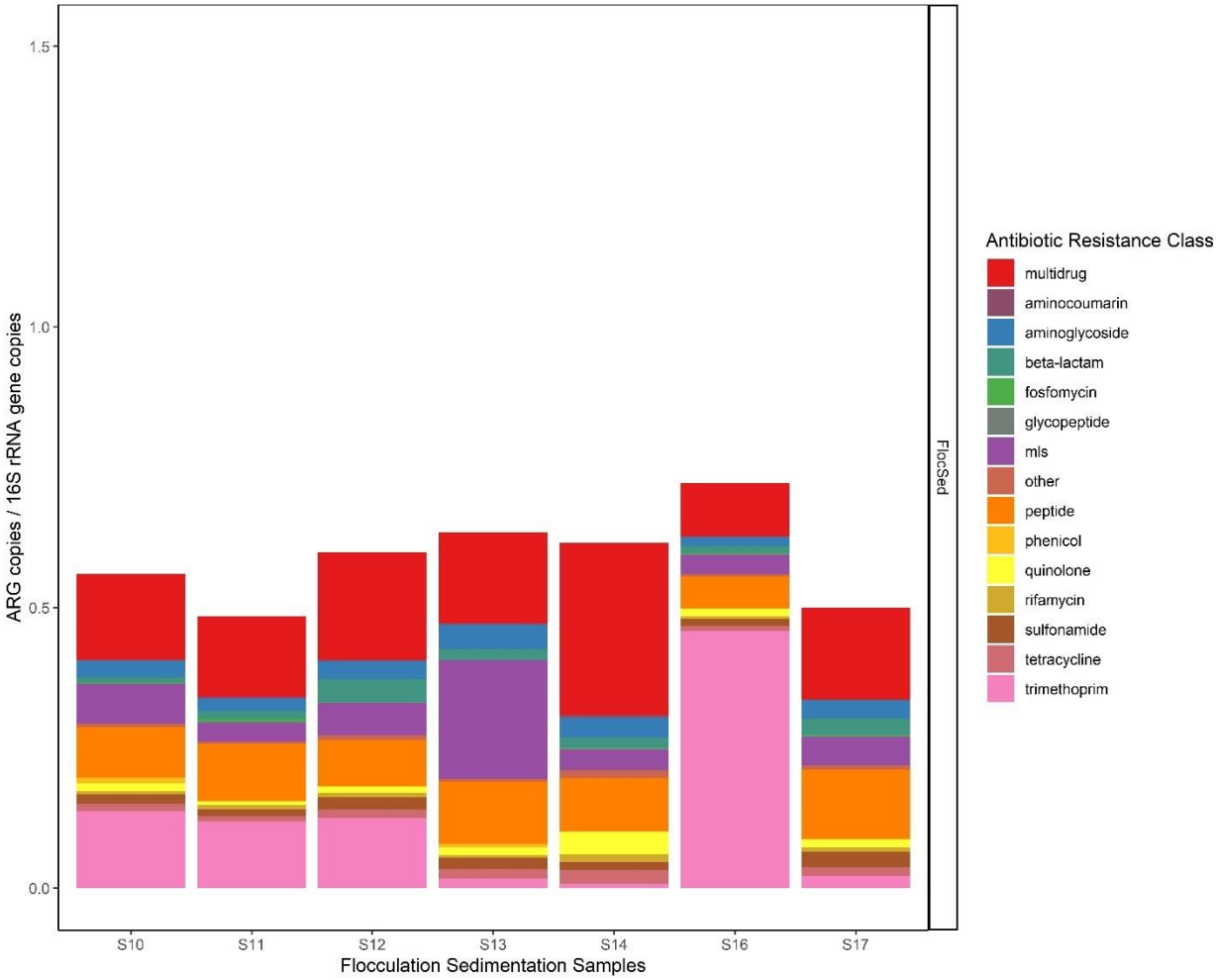
Subsection 1: Supplemental Figures associated with Trends in Classified ARG Abundance Detected throughout the Treatment Train



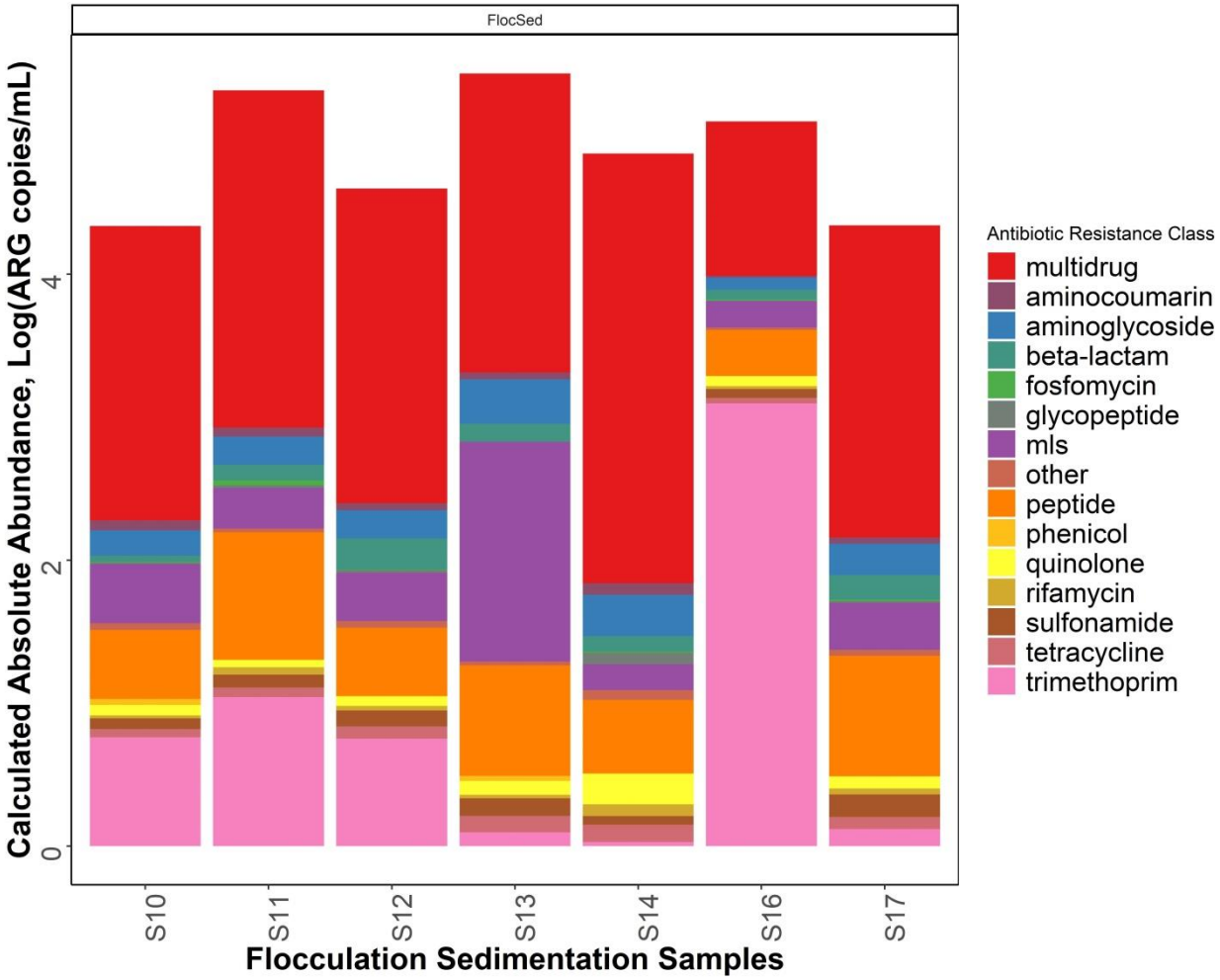
SI Figure D.10: Relative abundance (ARG gene copies normalized to 16S rRNA gene copies) of antibiotic resistance classes identified via metagenomic analysis of samples collected from Pilot Inlet at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.



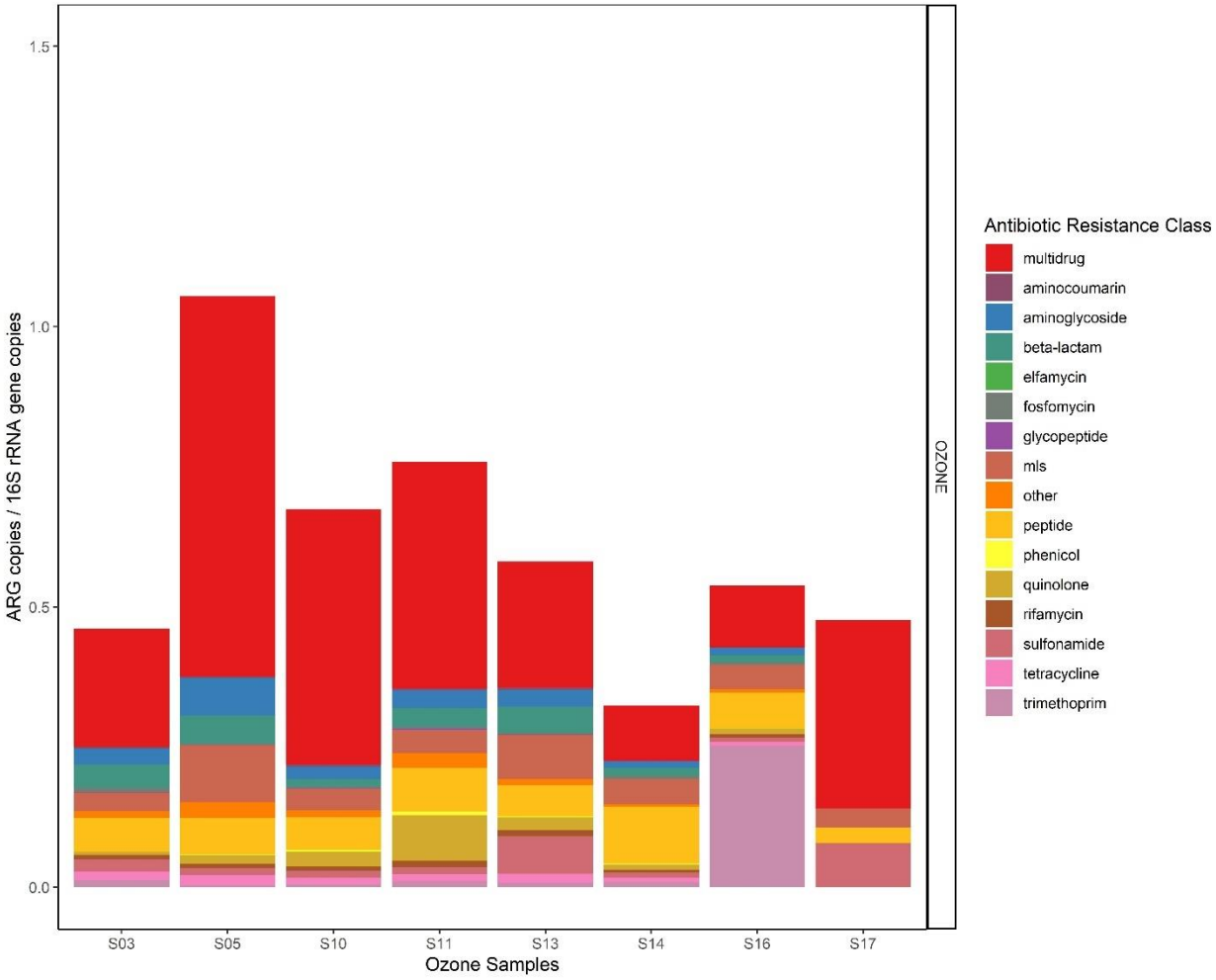
SI Figure D.11: Absolute abundance (log function of ARG gene copies per mL) of antibiotic resistance classes identified via metagenomic analysis of samples collected from Pilot Inlet at the SWIFT Pilot Facility at York River site between January 2017 and April 2018. Total abundances were calculated by multiplying relative abundance data from metagenomic sequencing by 16S rRNA gene copy numbers obtained by qPCR.



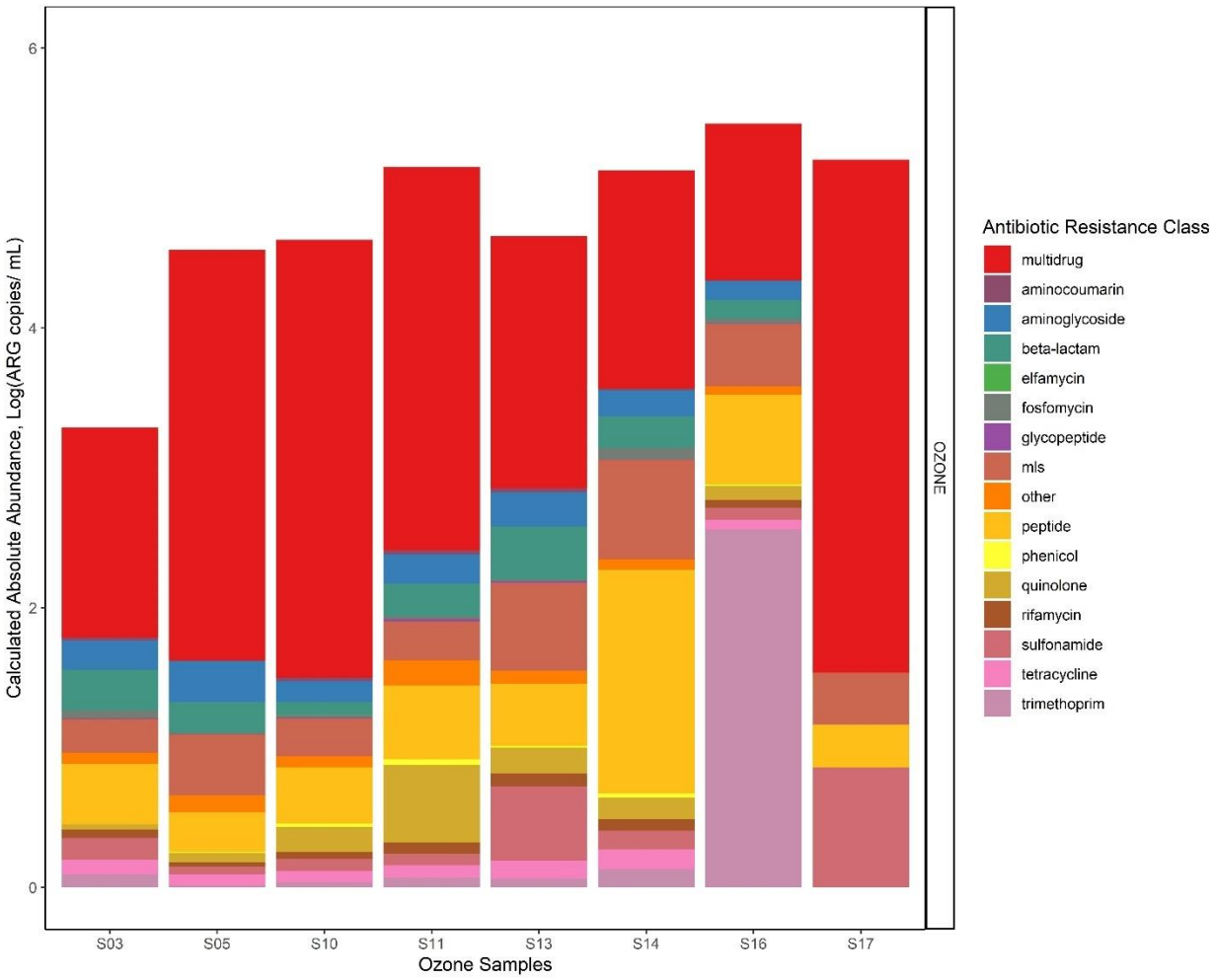
SI Figure D.12: Relative abundance (ARG gene copies normalized to 16S rRNA gene copies) of antibiotic resistance classes identified via metagenomic analysis of samples collected from Flocculation-Sedimentation at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.



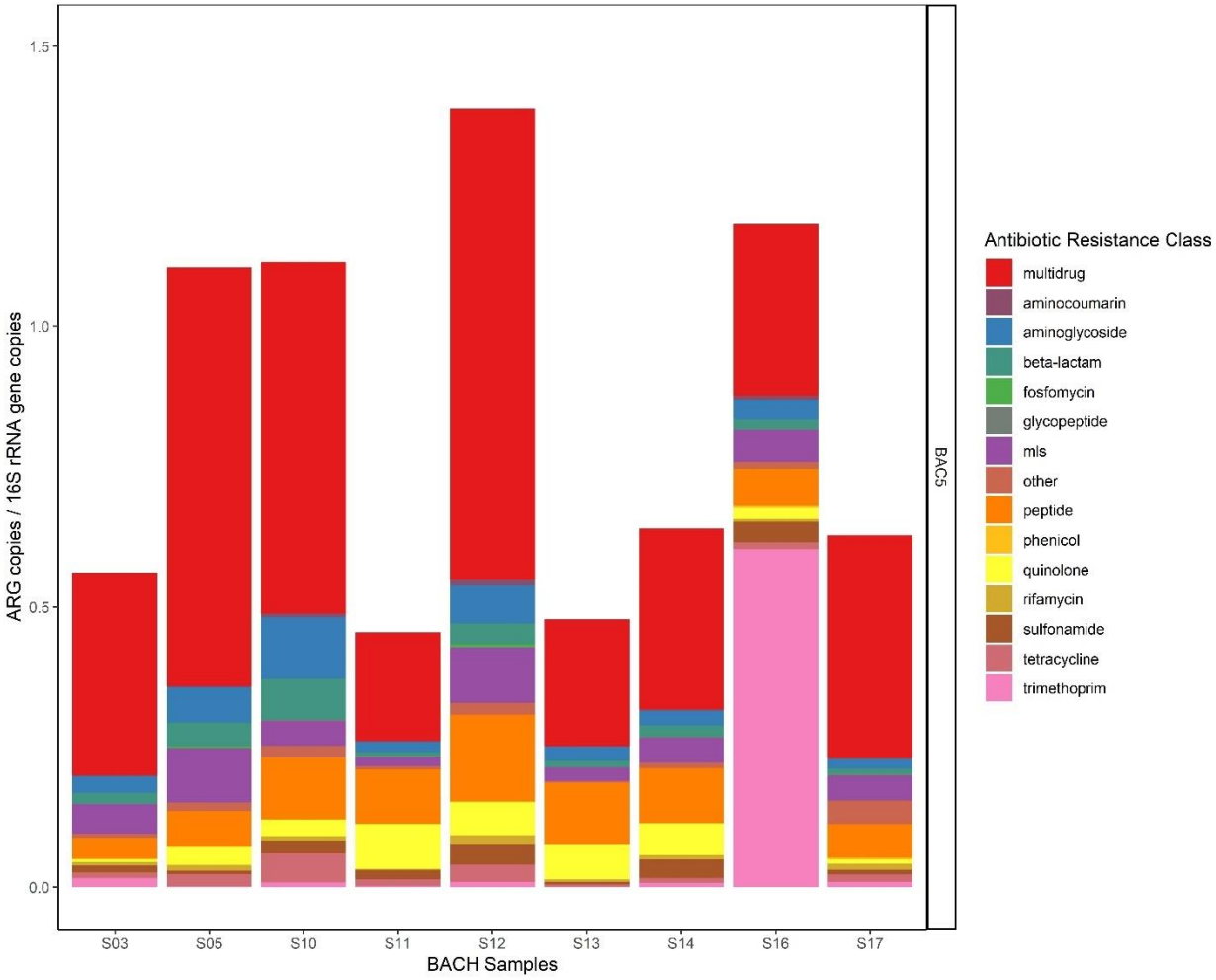
SI Figure D.13: Absolute abundance of (log function of ARG gene copies per mL) of antibiotic resistance classes identified via metagenomic analysis of samples collected from Flocculation-Sedimentation at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.



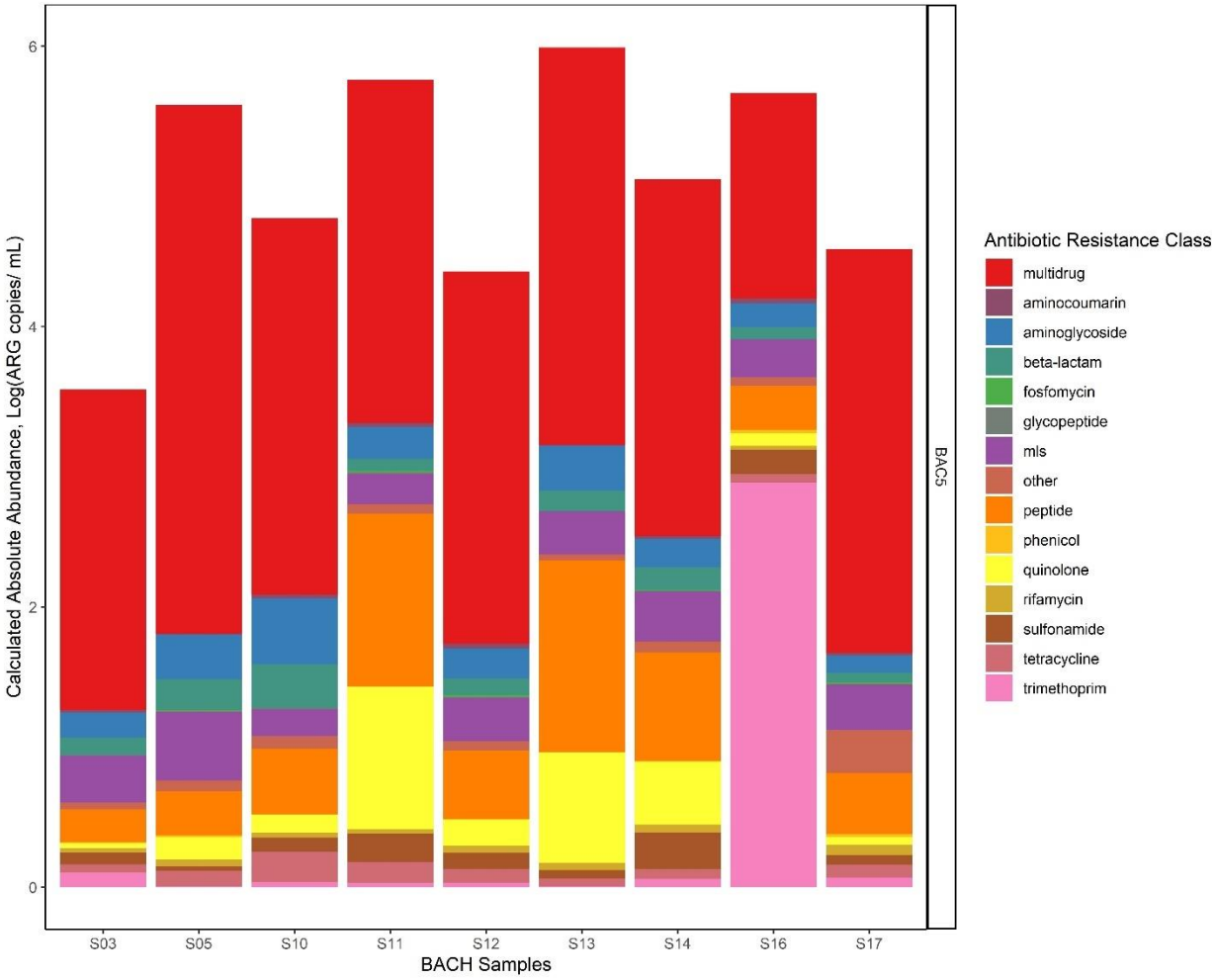
SI Figure D.14: **Relative abundance (ARG gene copies normalized to 16S rRNA gene copies) of antibiotic resistance classes identified via metagenomic analysis of samples collected from Ozone effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.**



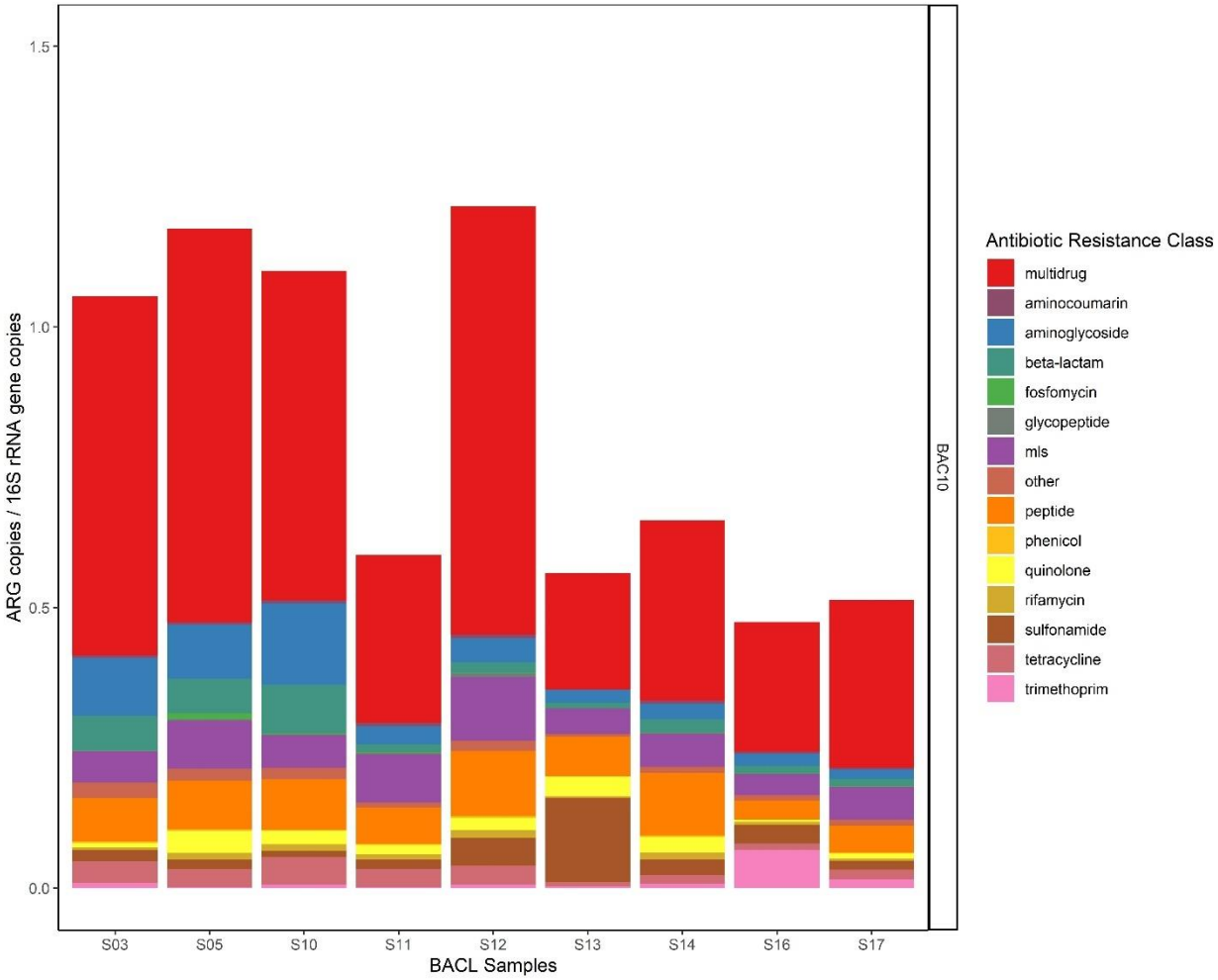
SI Figure D.15: Absolute abundance (log function of ARG gene copies per mL) of antibiotic resistance classes identified via metagenomic analysis of samples collected from Ozone effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.



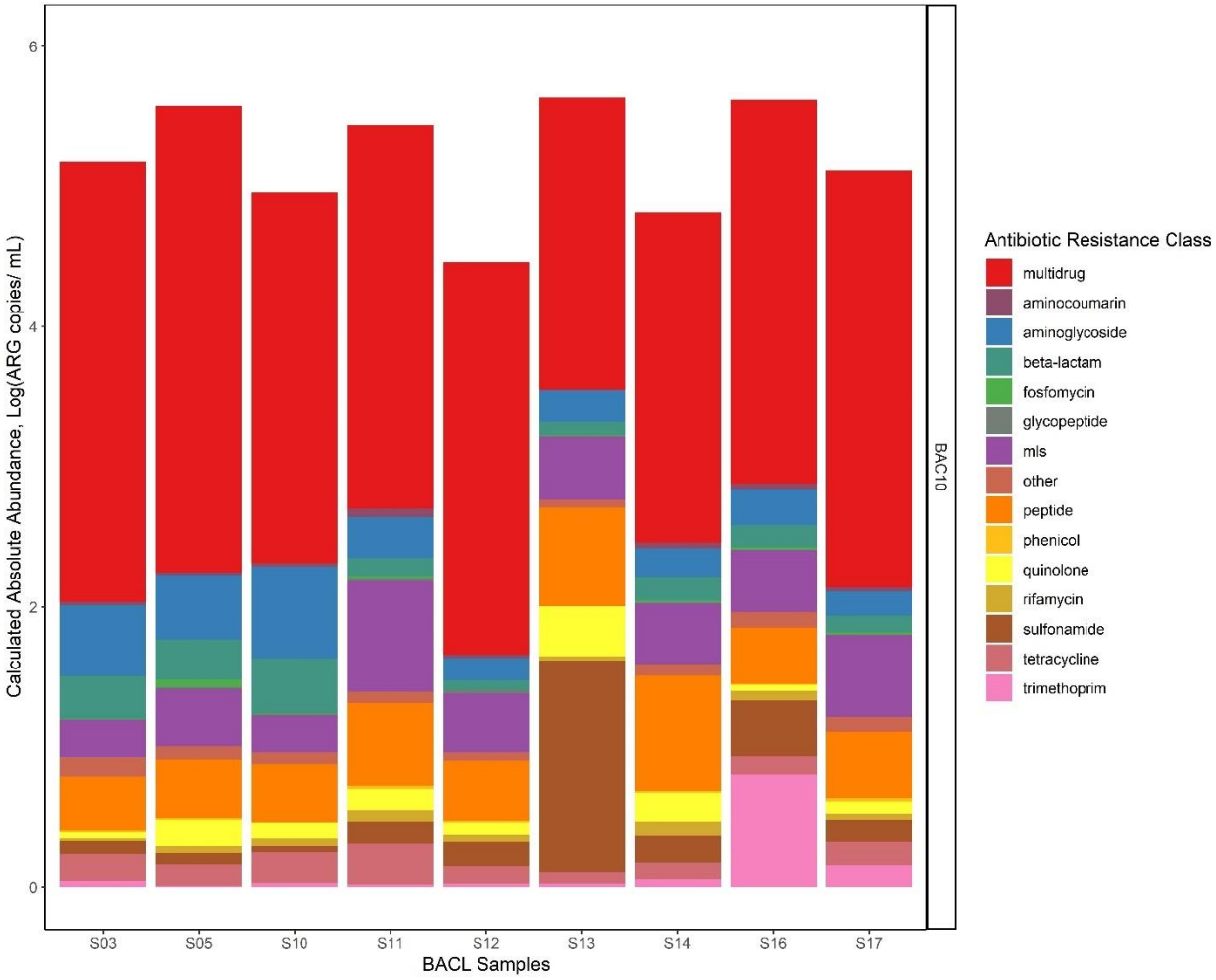
SI Figure D.16: Relative abundance (ARG gene copies normalized to 16S rRNA gene copies) of antibiotic resistance classes identified via metagenomic analysis of samples collected from BAC-5 effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.



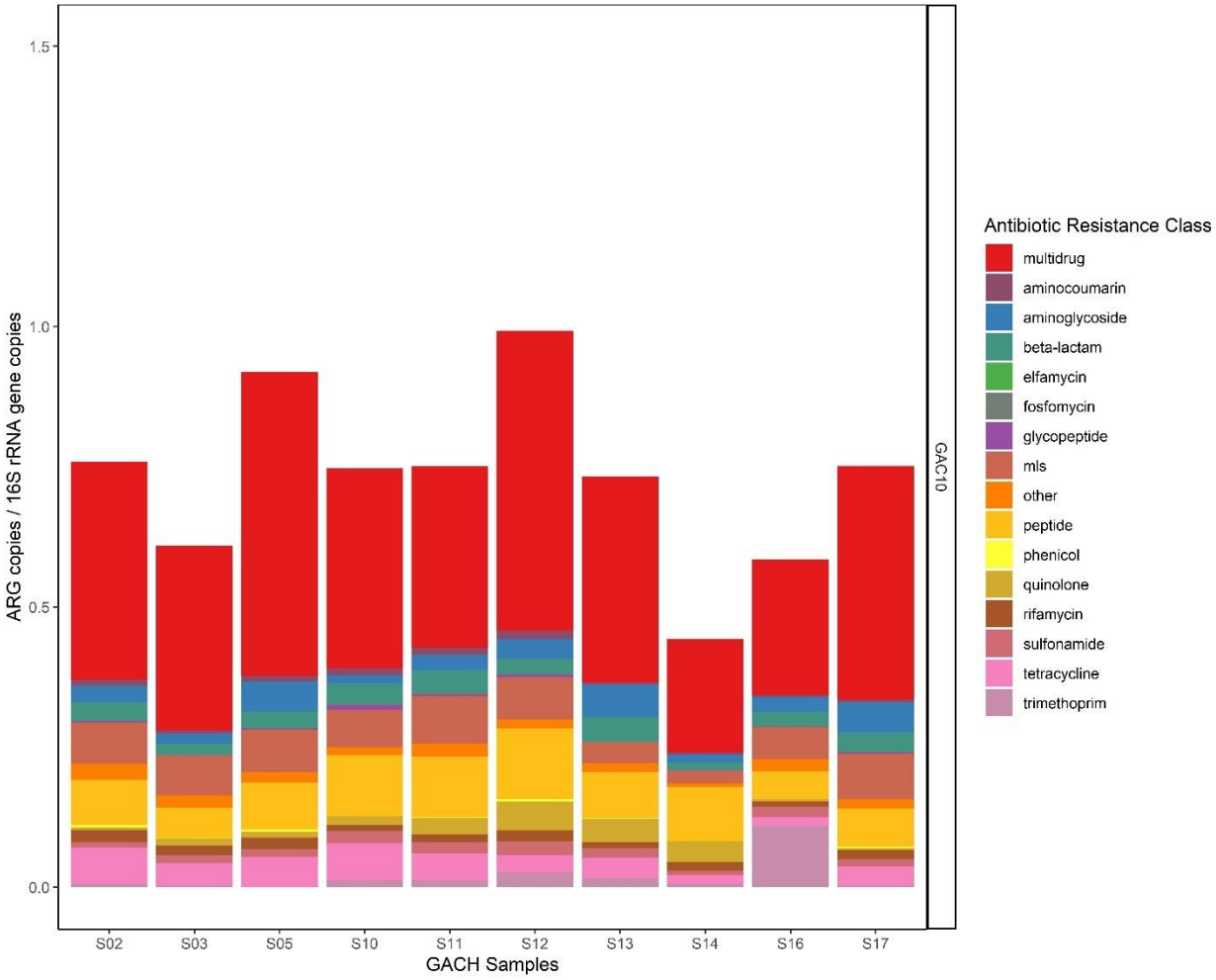
SI Figure D.17: Absolute abundance (log function of ARG gene copies per mL) of antibiotic resistance classes identified via metagenomic analysis of samples collected from BAC-5 effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.



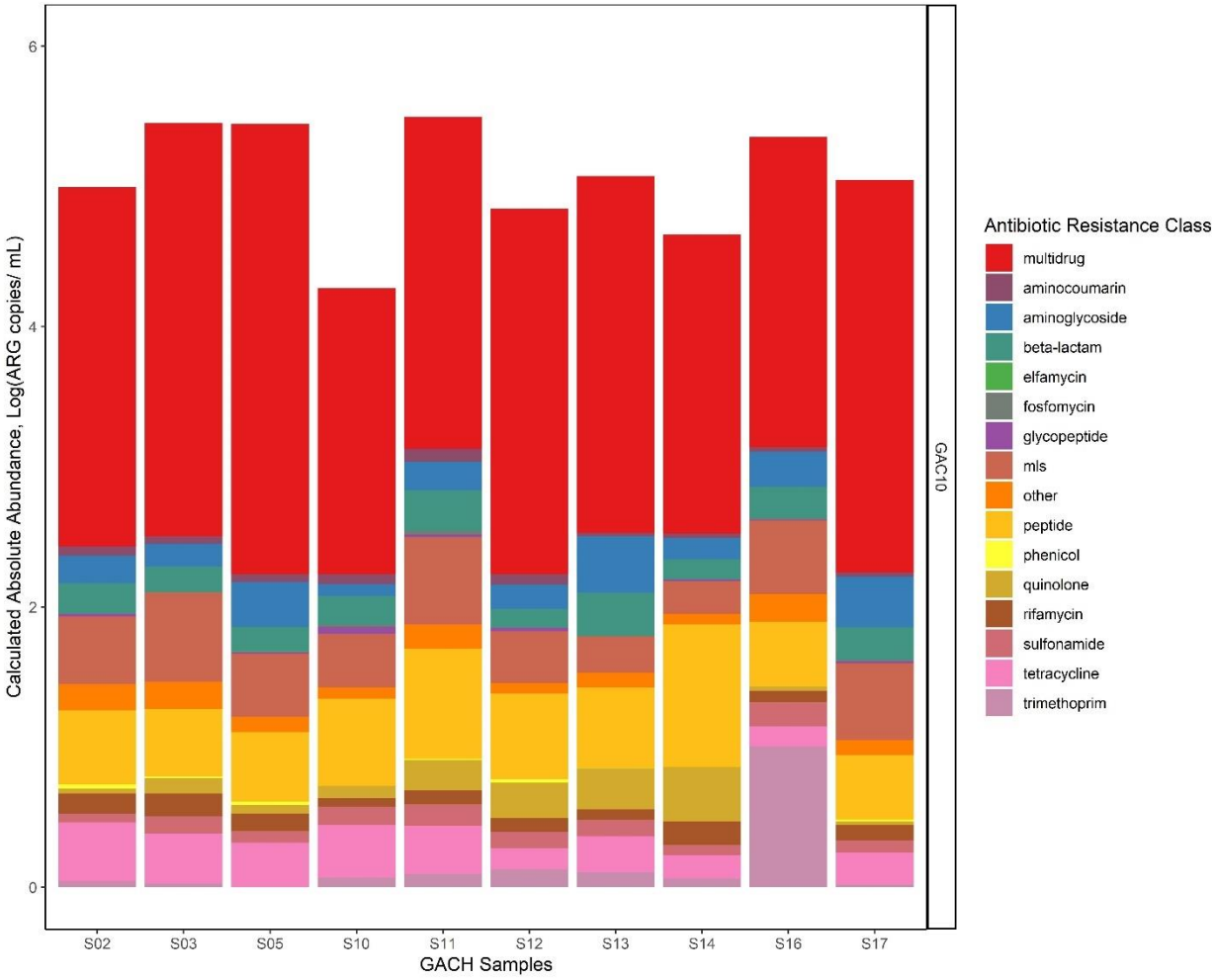
SI Figure D.18: Relative abundance (ARG gene copies normalized to 16S rRNA gene copies) of antibiotic resistance classes identified via metagenomic analysis of samples collected from BAC-10 effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.



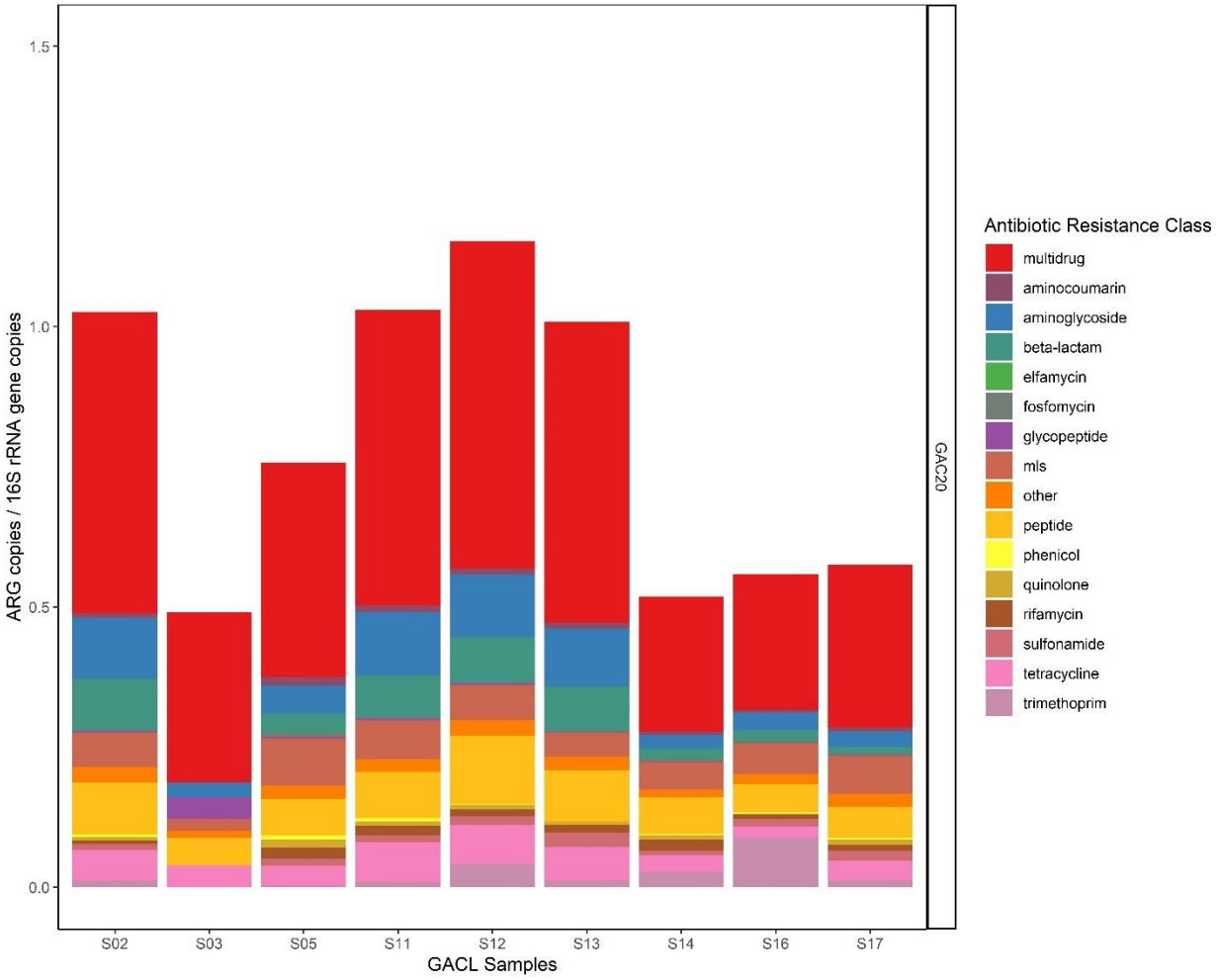
SI Figure D.19: Absolute abundance (log function of ARG gene copies per mL) of antibiotic resistance classes identified via metagenomic analysis of samples collected from BAC-10 at the



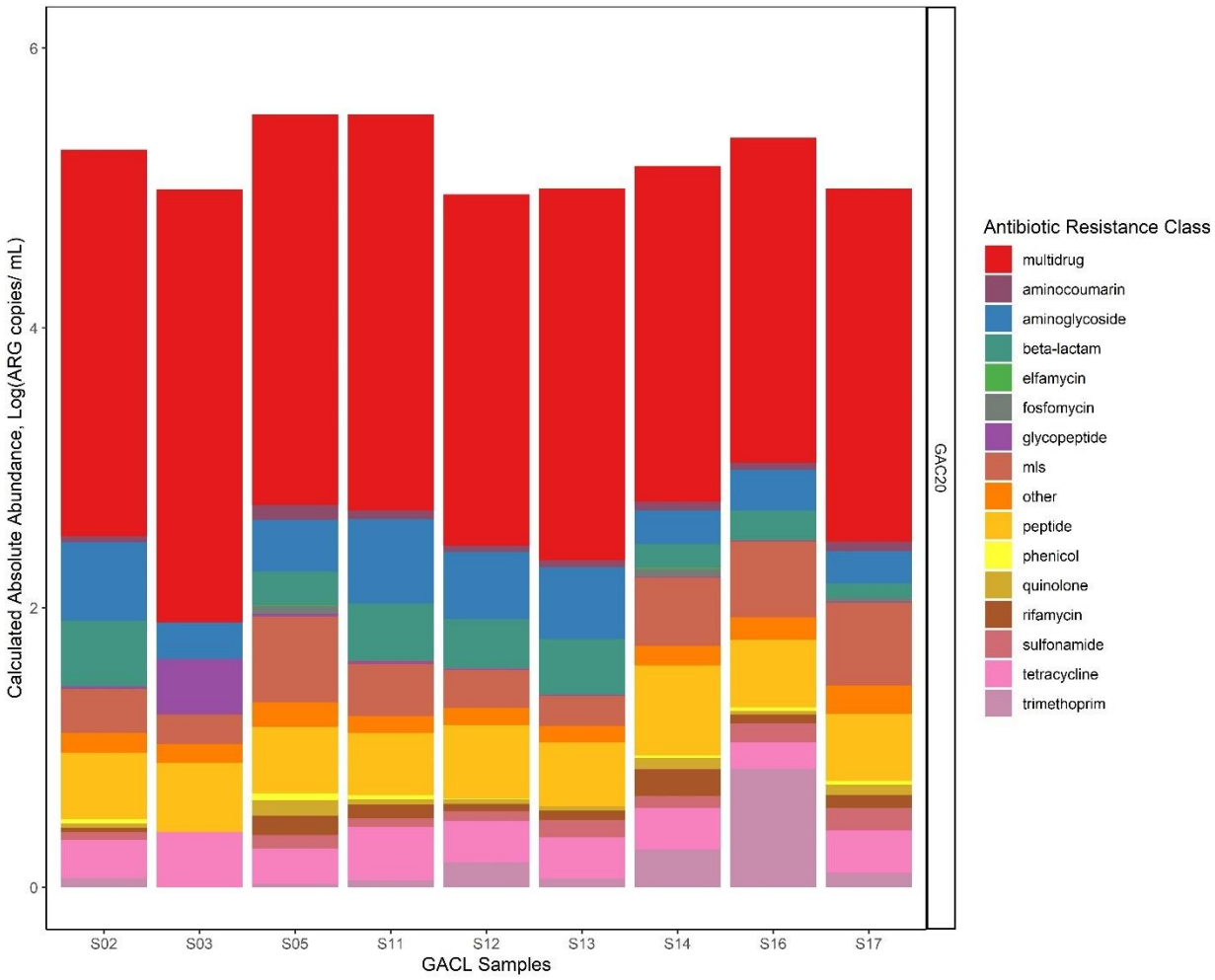
SI Figure D.20: Relative abundance (ARG gene copies normalized to 16S rRNA gene copies) of antibiotic resistance classes identified via metagenomic analysis of samples collected from GAC-10 effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.



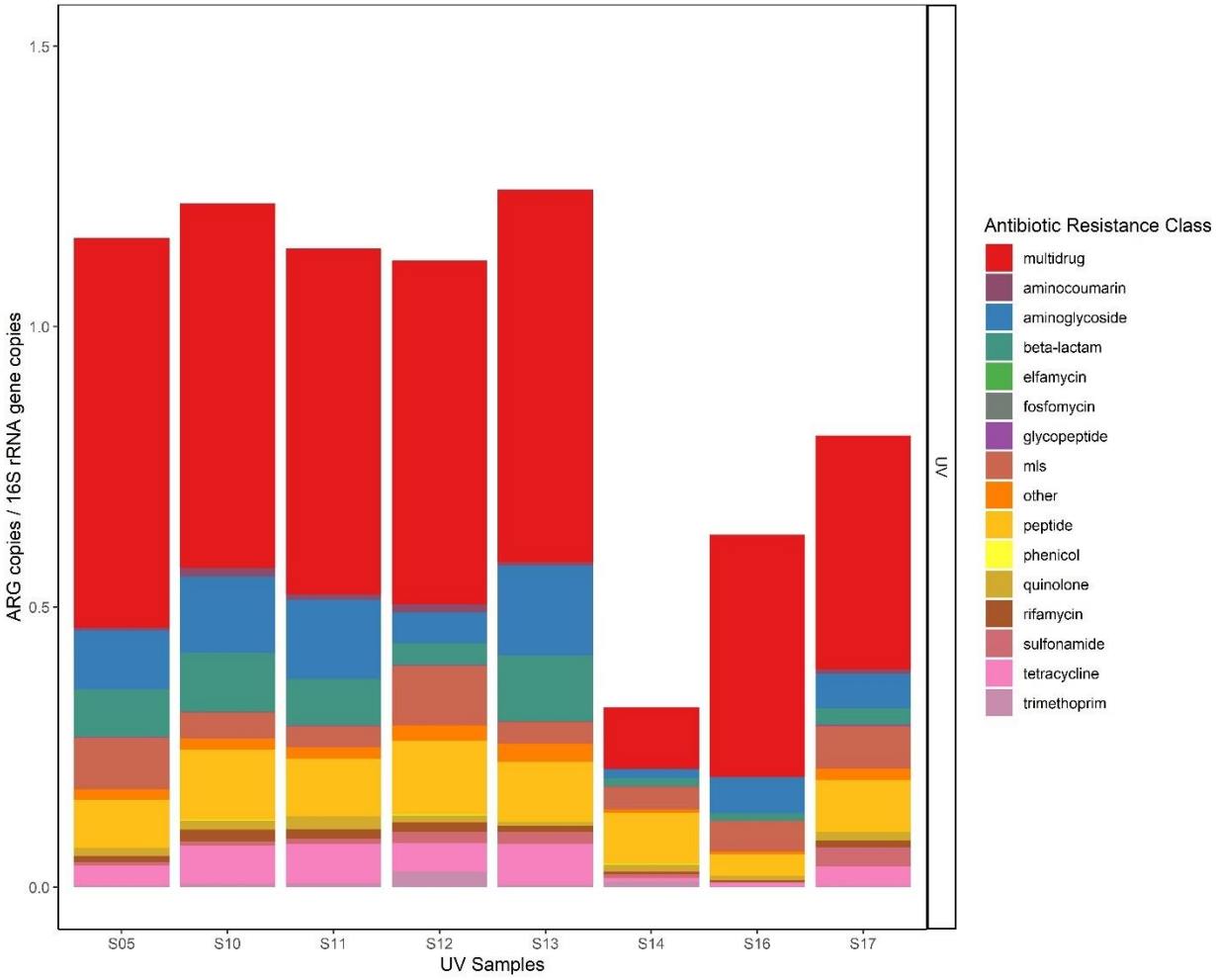
SI Figure D.21: Absolute abundance (log function of ARG gene copies per mL) of antibiotic resistance classes identified via metagenomic analysis of samples collected from GAC-10 effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.



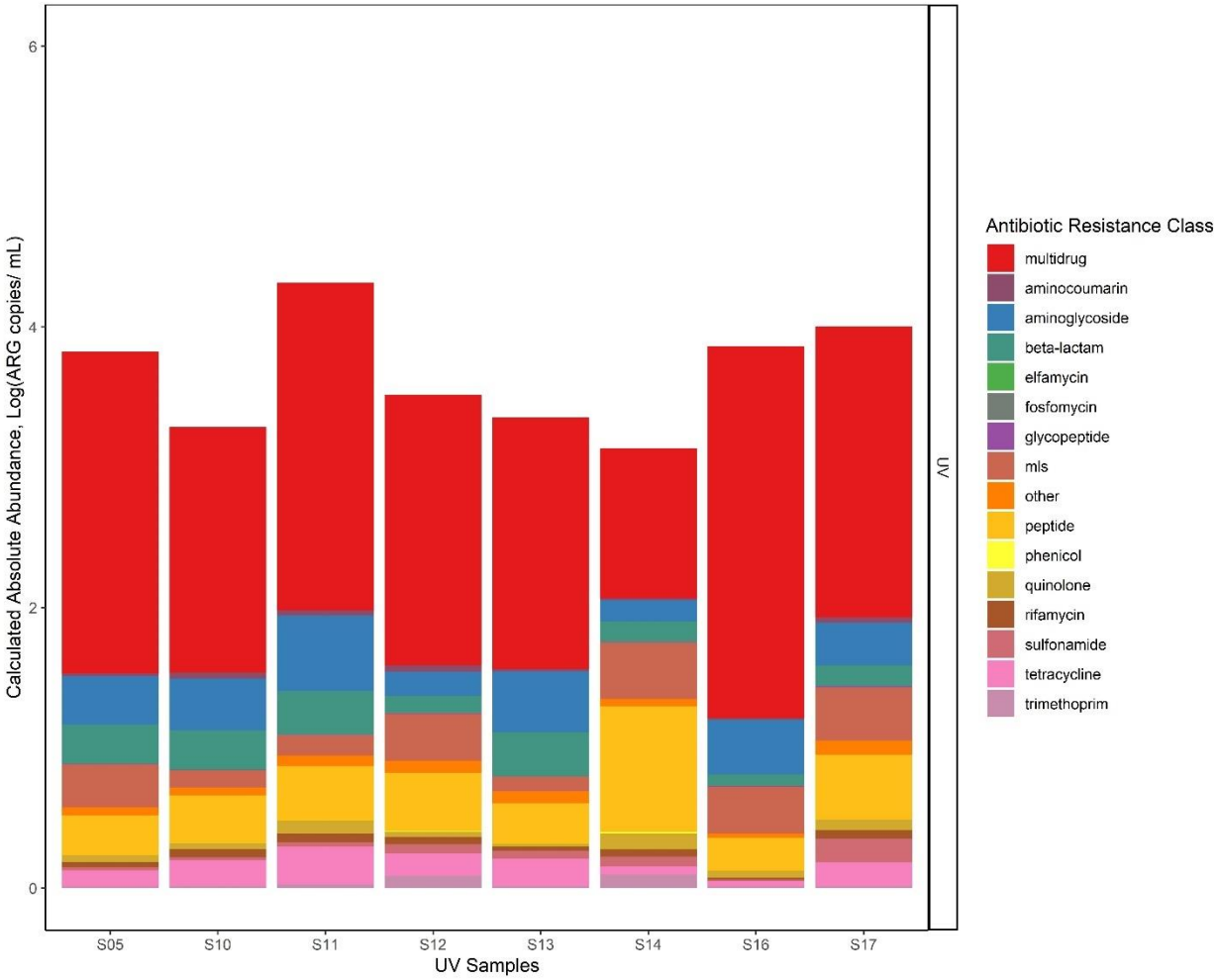
SI Figure D.22: Relative abundance (ARG gene copies normalized to 16S rRNA gene copies) of antibiotic resistance classes identified via metagenomic analysis of samples collected from GAC-20 effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.



SI Figure D.23: Absolute abundance (log function of ARG gene copies per mL) of antibiotic resistance classes identified via metagenomic analysis of samples collected from GAC-20 effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.

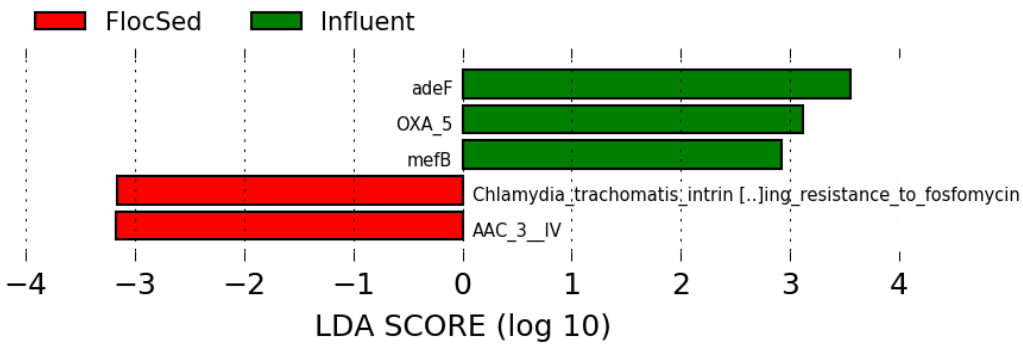


SI Figure D.24: Relative abundance (ARG gene copies normalized to 16S rRNA gene copies) of antibiotic resistance classes identified via metagenomic analysis of samples collected from UV effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.

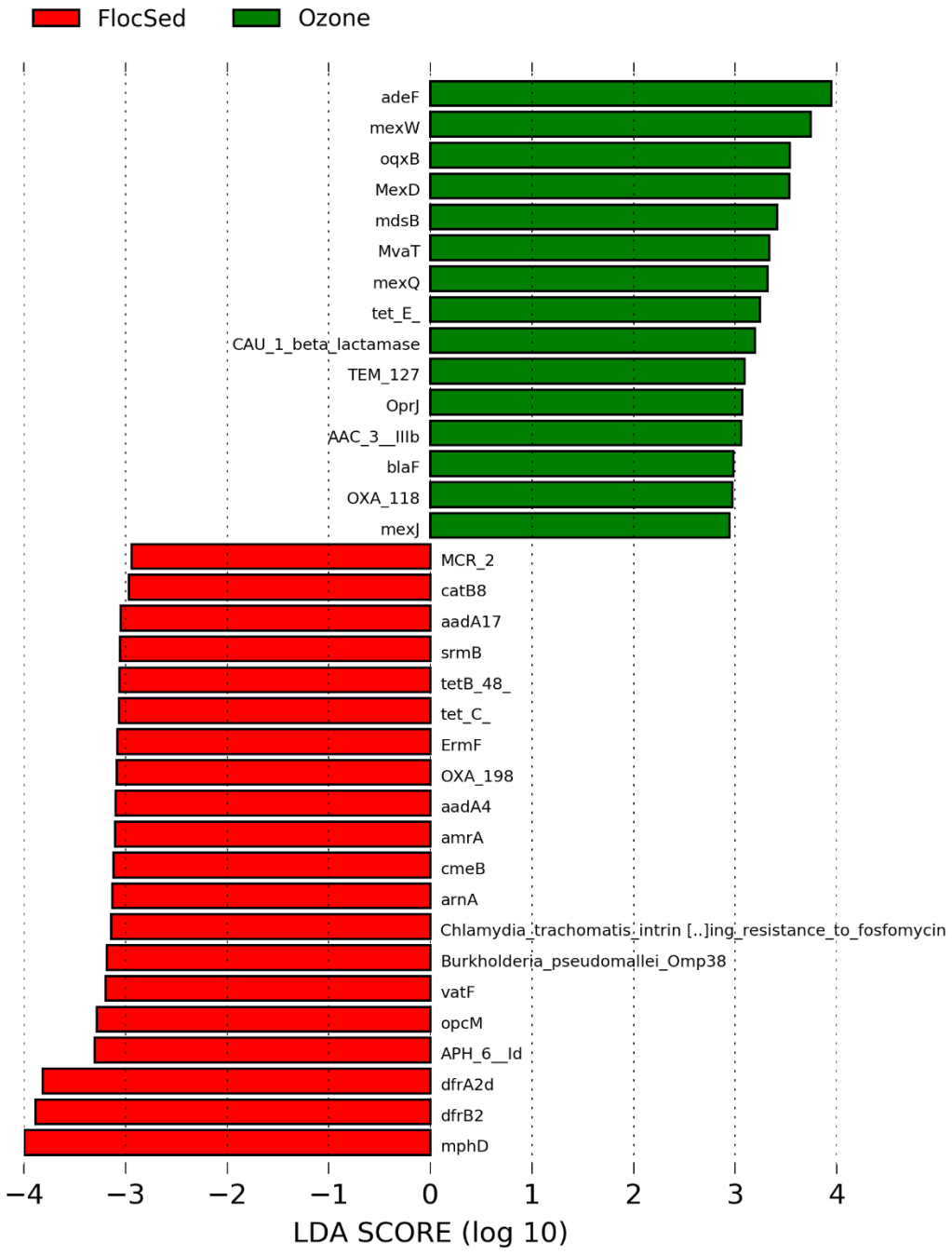


SI Figure D.25: Absolute abundance (log function of ARG gene copies per mL) of antibiotic resistance classes identified via metagenomic analysis of samples collected from UV effluent at the SWIFT Pilot Facility at York River site between January 2017 and April 2018.

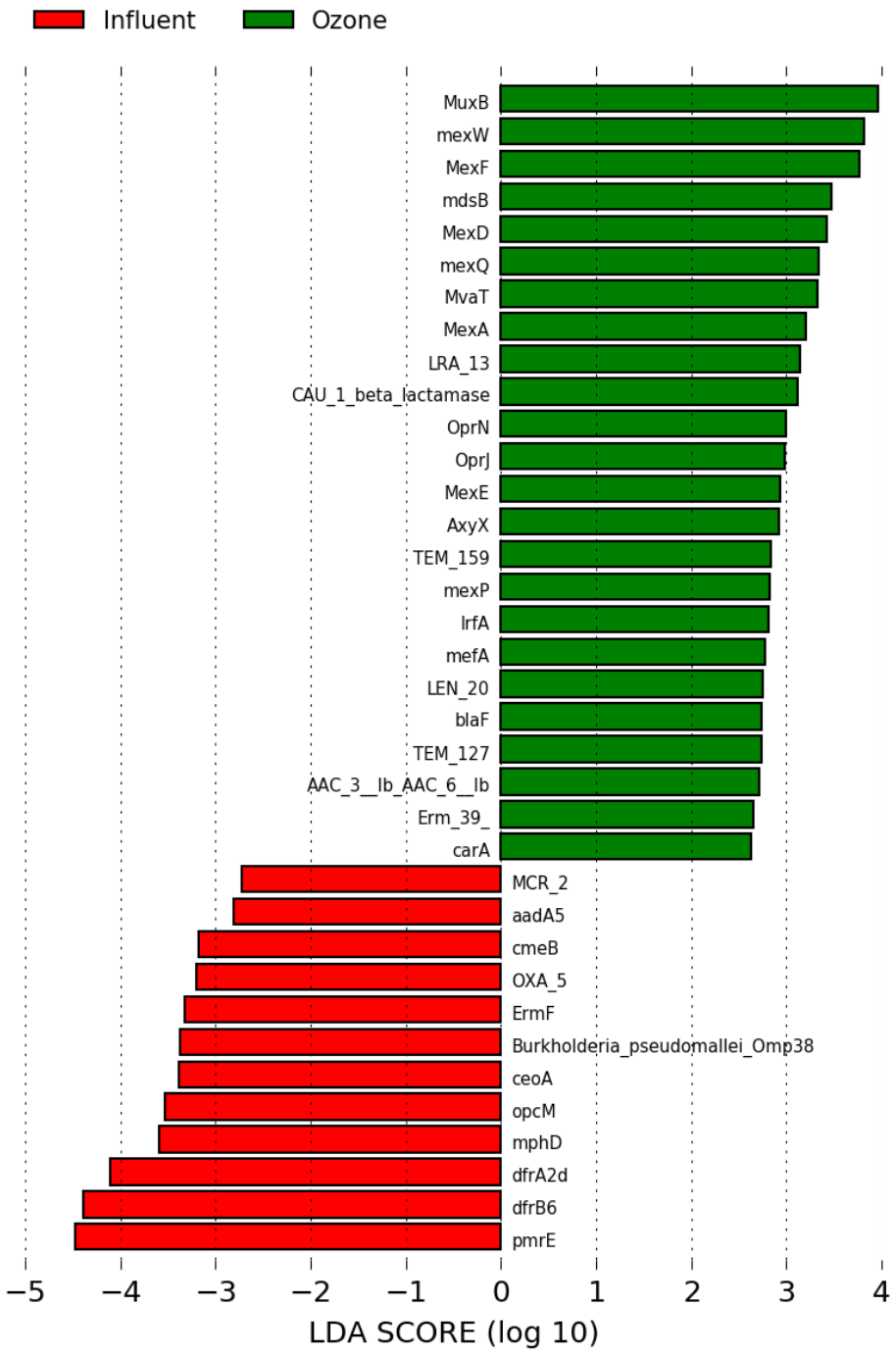
Subsection 2: Supplemental Figures associated with Identification of Discriminatorily Abundant ARGs



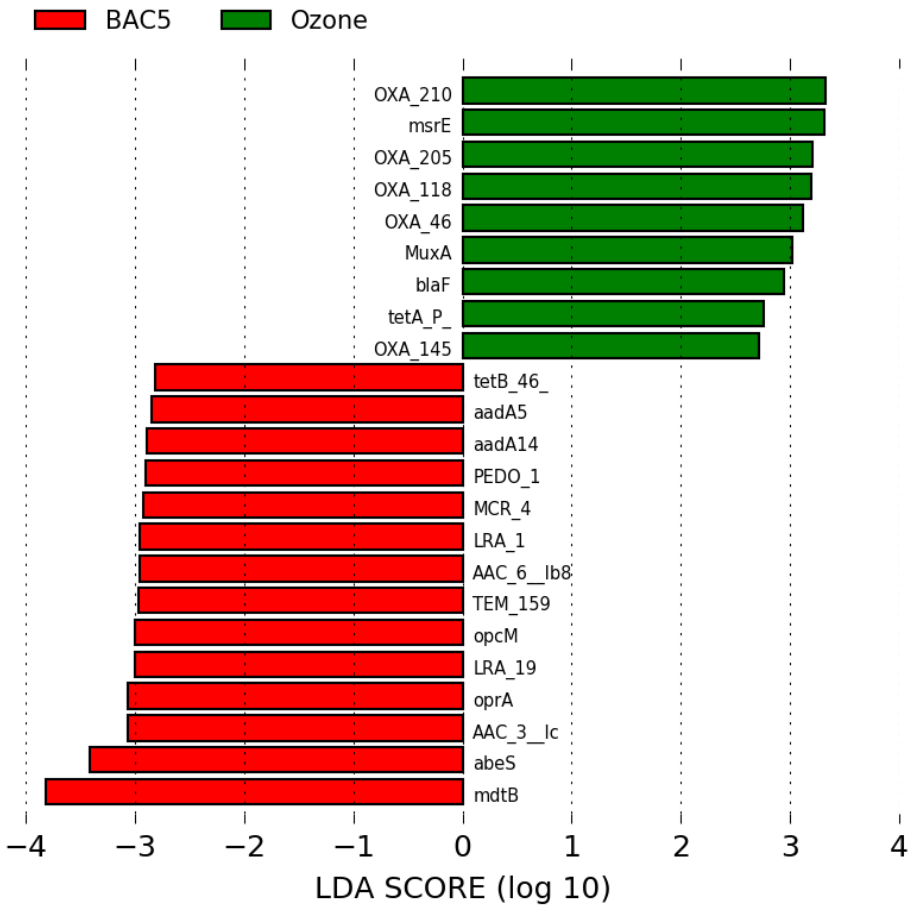
SI Figure D.26: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for influent and flocsed effluent.



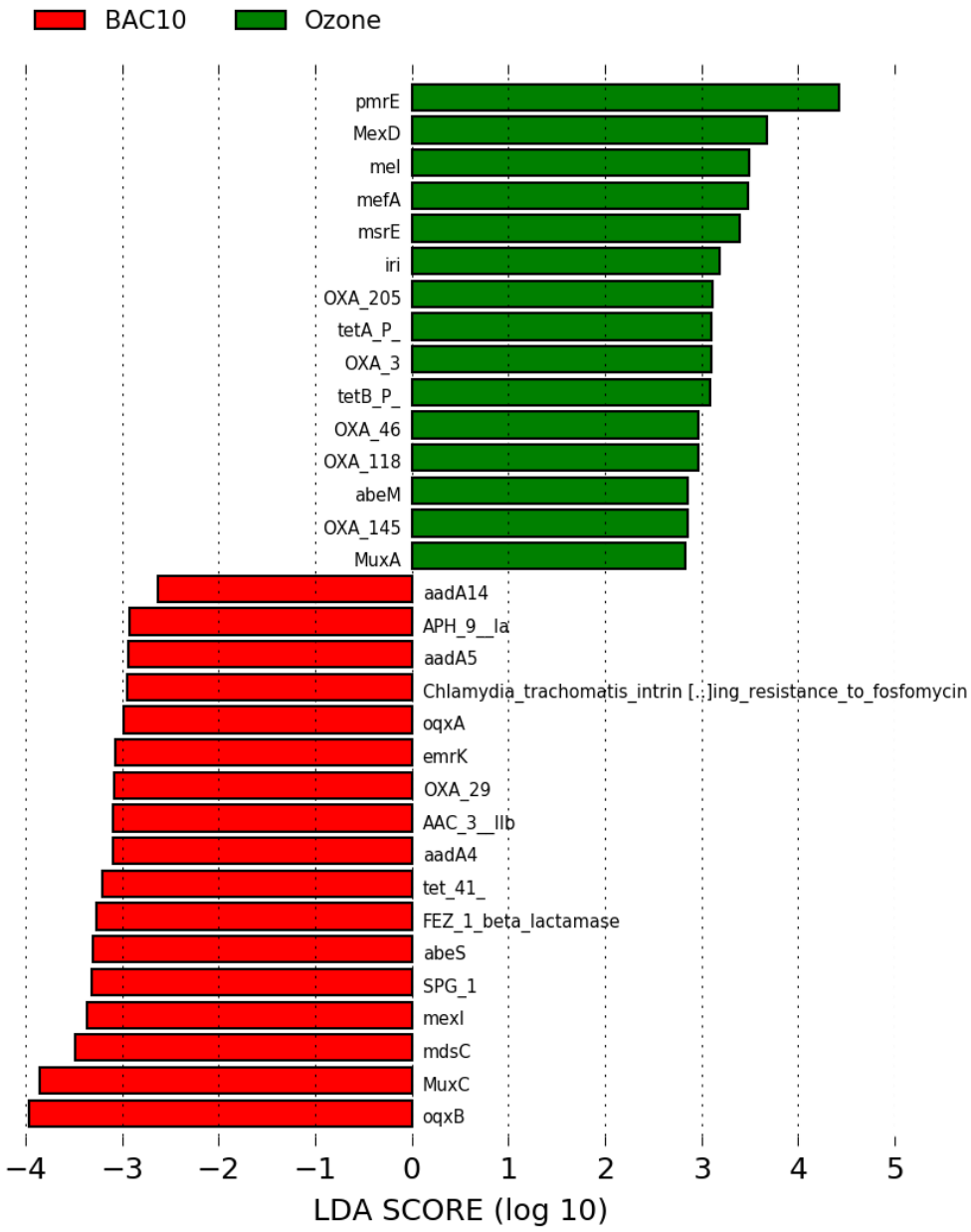
SI Figure D.27: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for flocsed and ozone effluent.



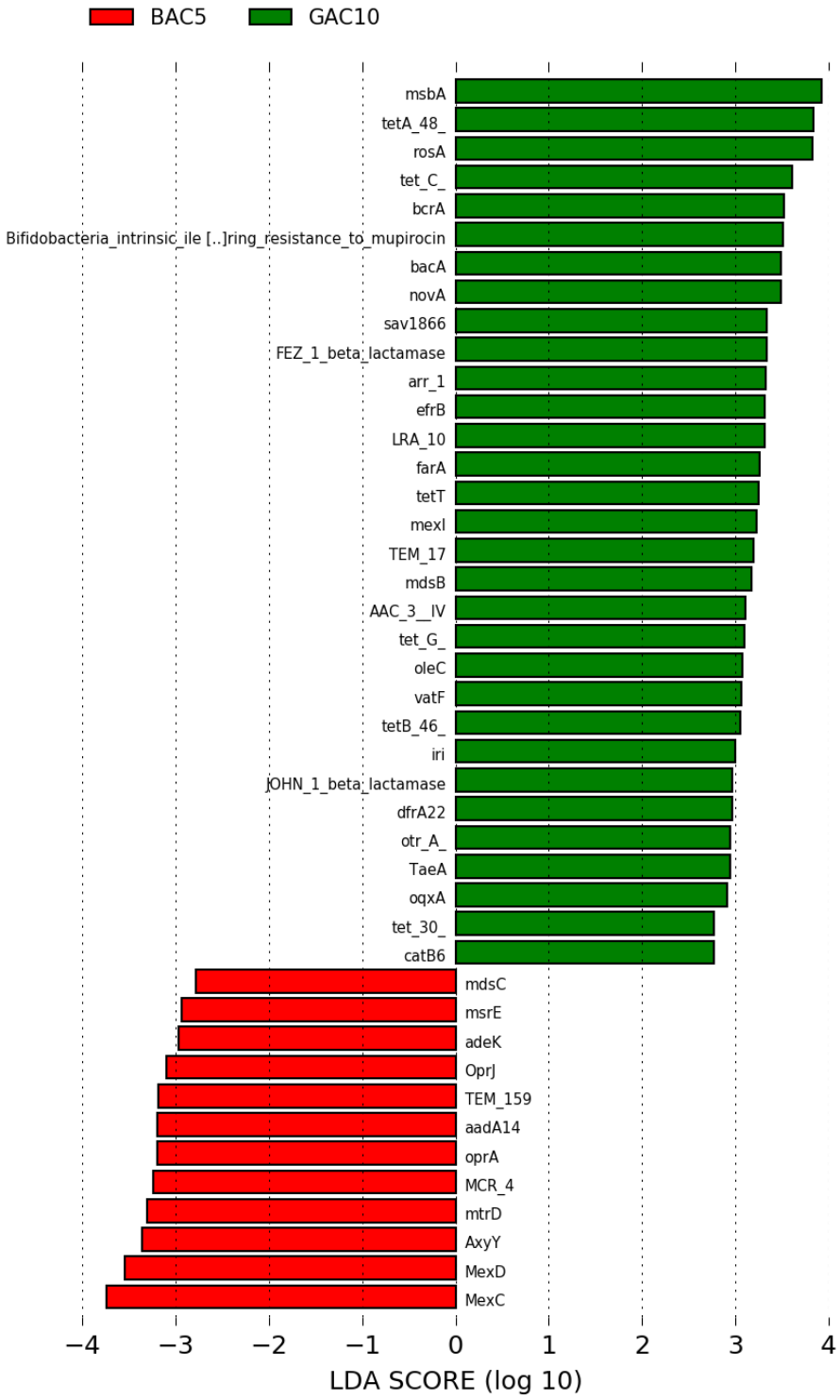
SI Figure D.28: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for influent and ozone effluent.



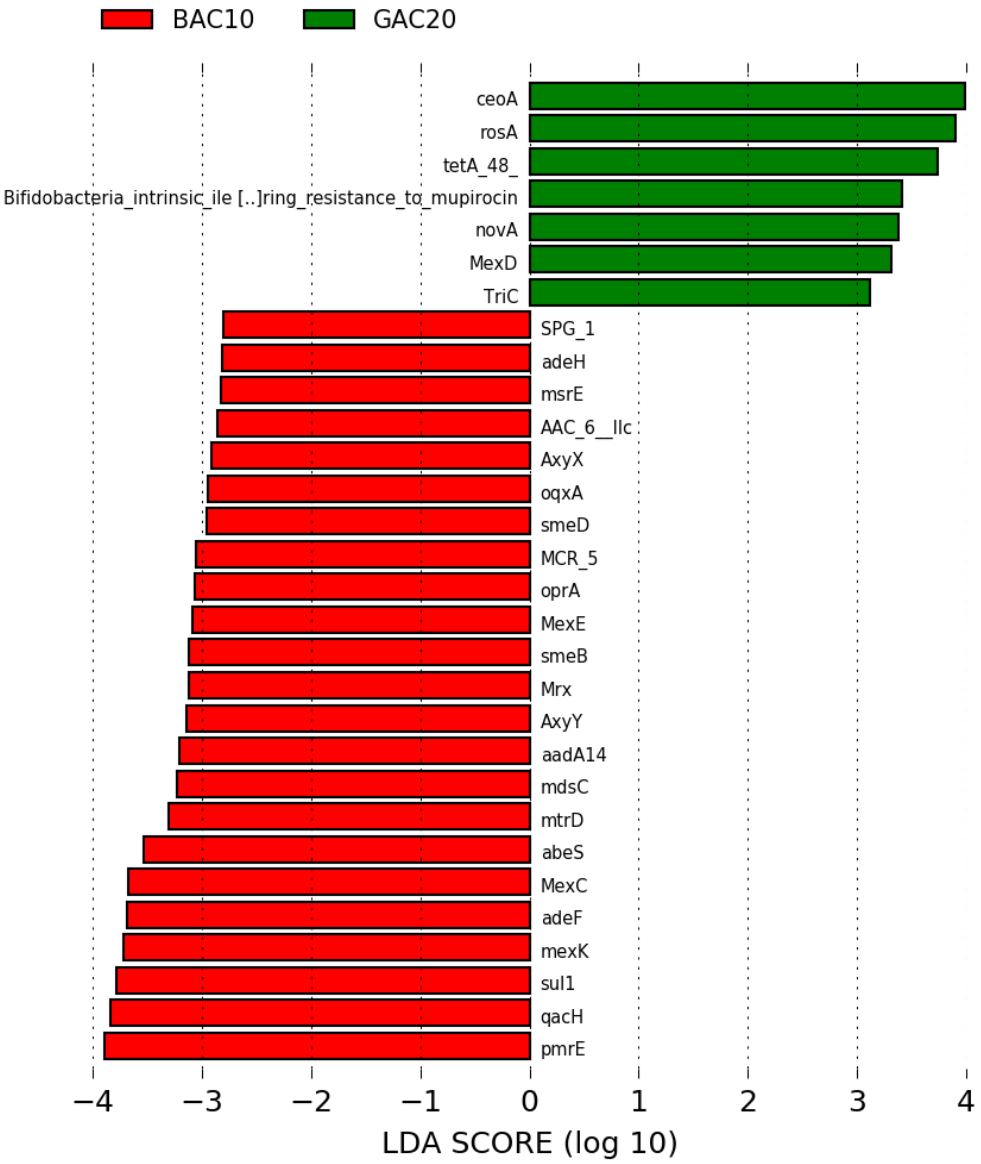
SI Figure D.29: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for ozone and BAC5 effluents.



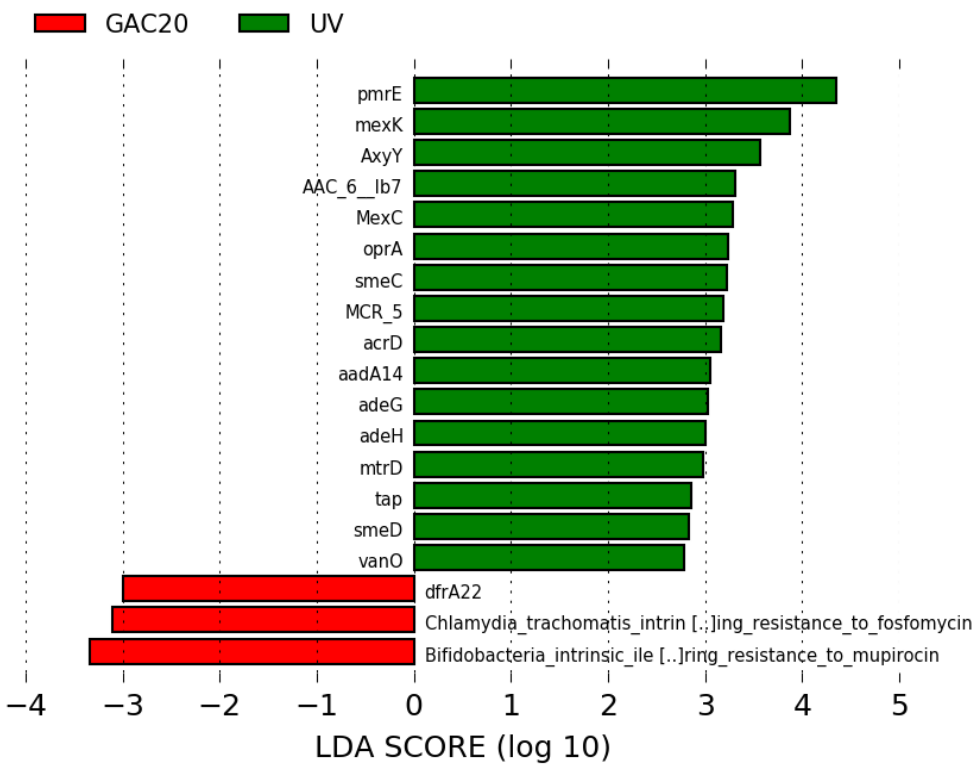
SI Figure D.30: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for ozone and BAC10 effluents.



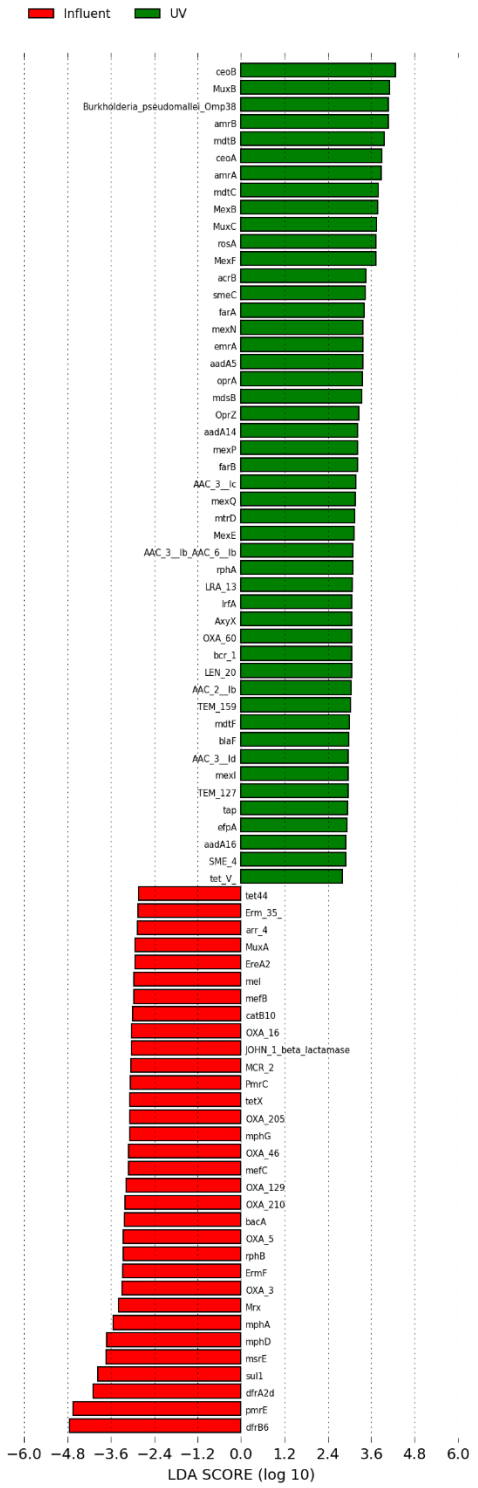
SI Figure D.31: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for BAC5 and GAC10.



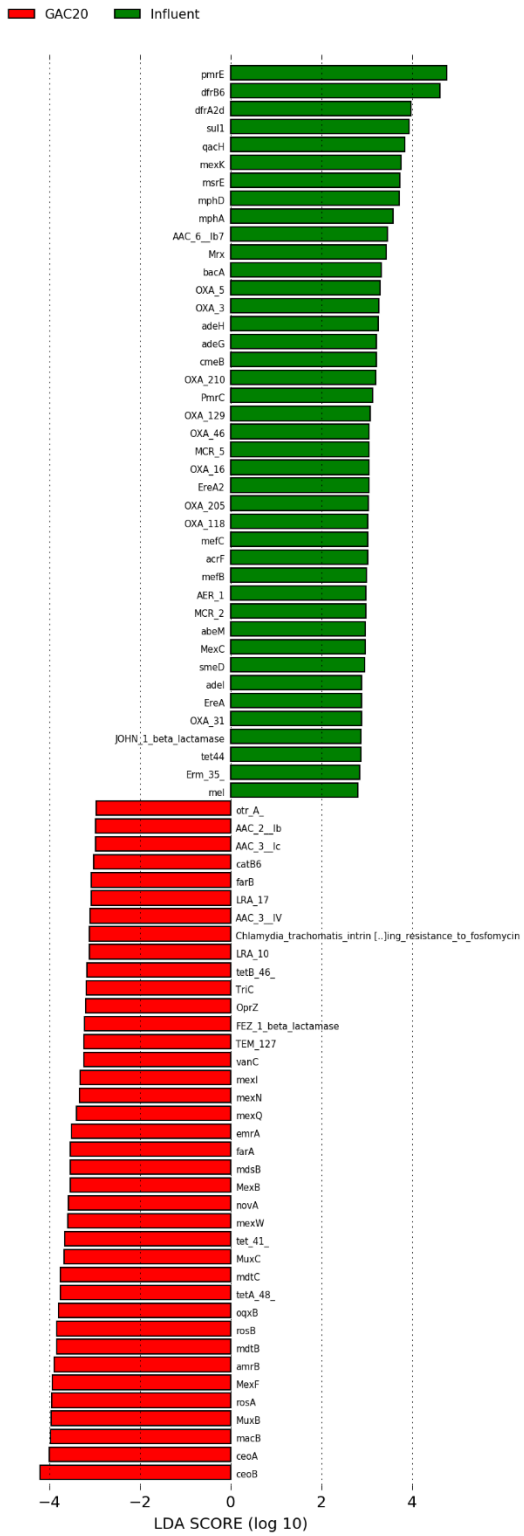
SI Figure D.32: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for BAC10 and GAC20.



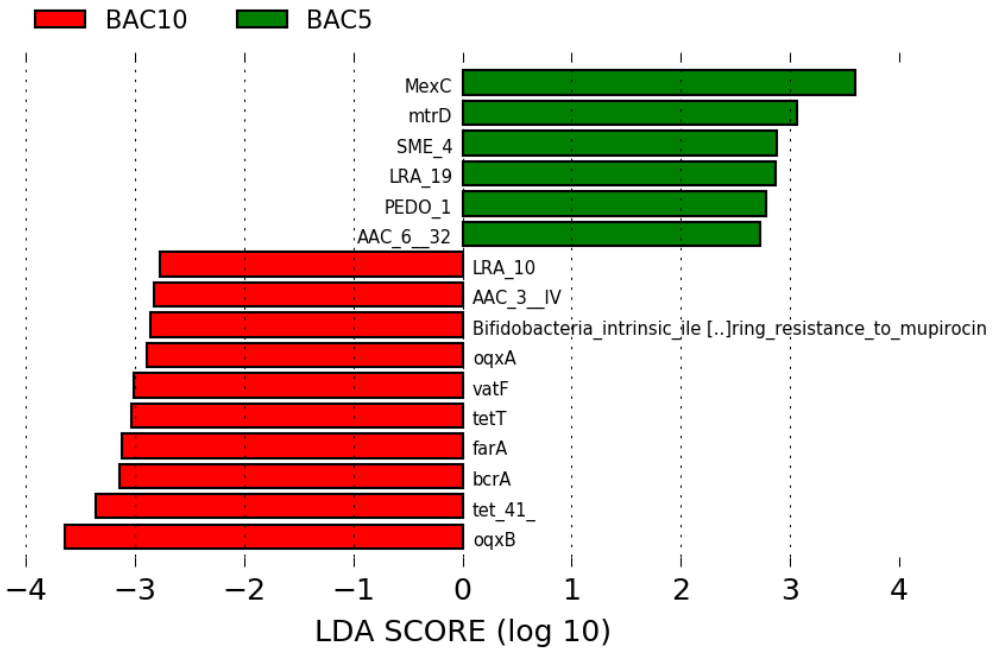
SI Figure D.33: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for GAC20 and UV effluents.



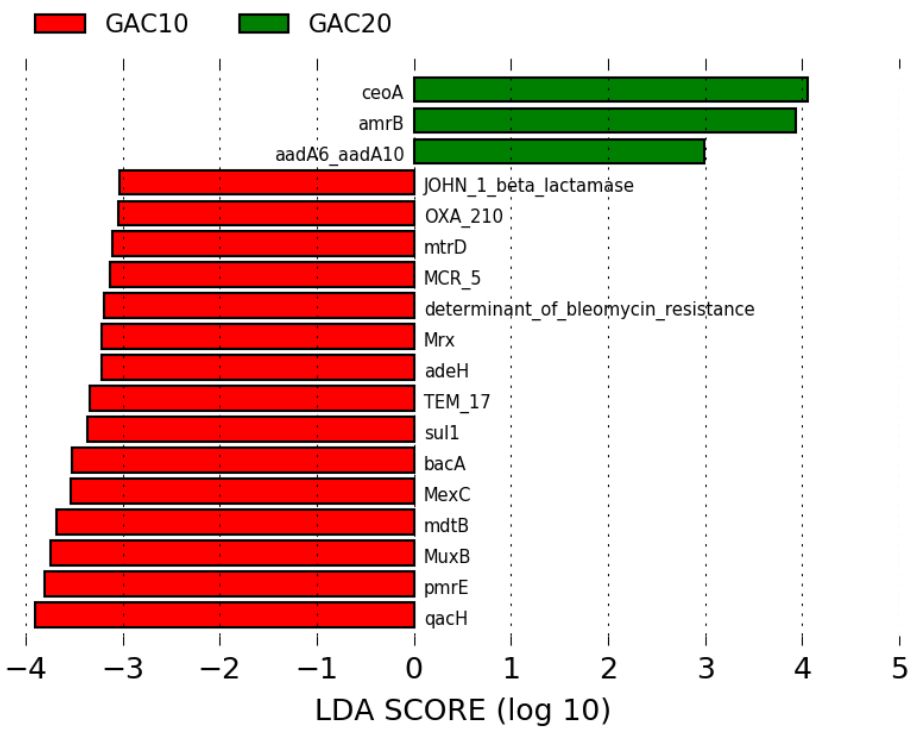
SI Figure D.34: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for influent and UV effluent.



SI Figure D.35: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for influent and GAC20 effluent.



SI Figure D.36: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for BAC5 and BAC10.



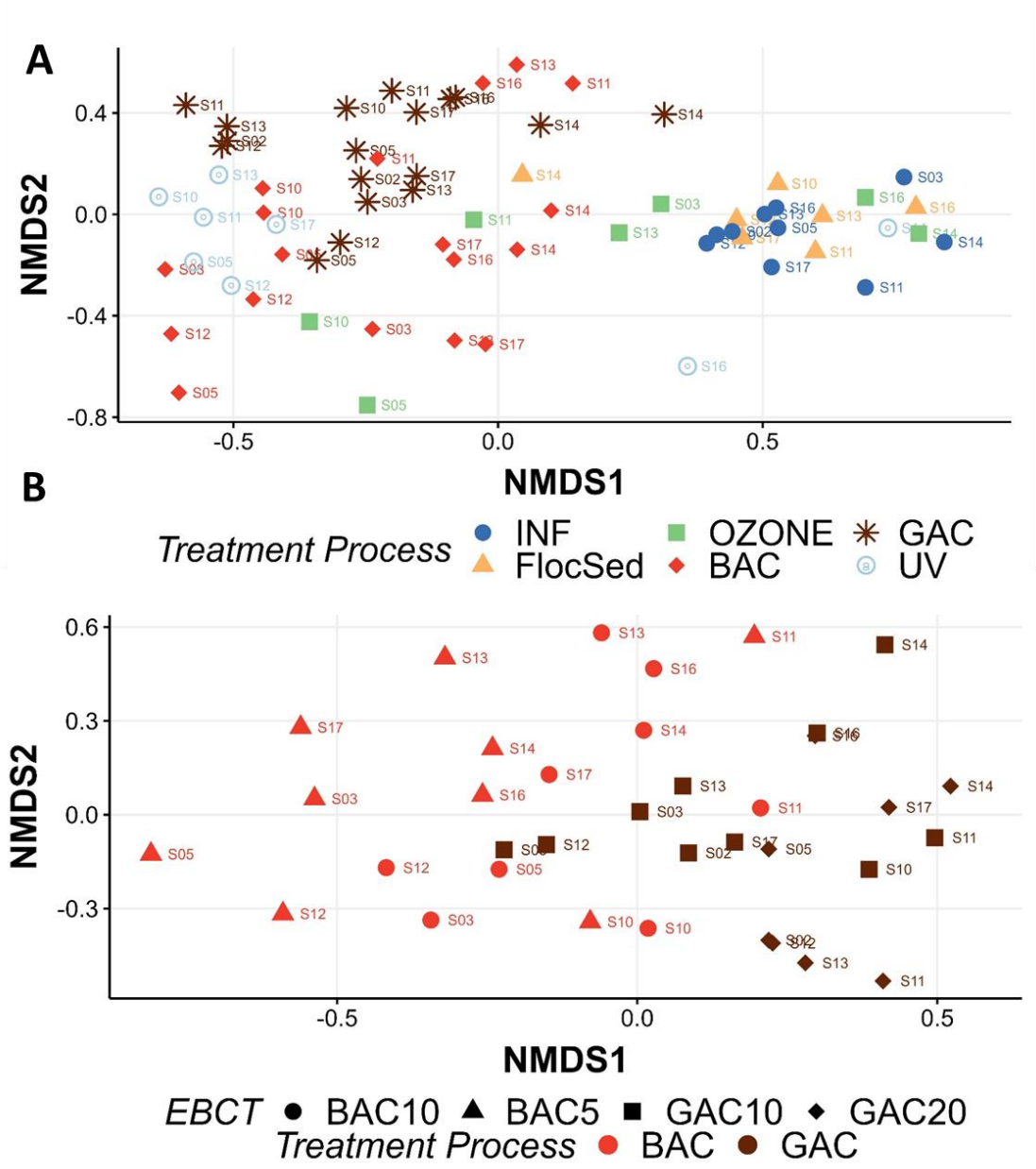
SI Figure D.37: Graphically presents the discriminatorily abundant ARGs (determined via LefSe analysis) for GAC10 and GAC20.

APPENDIX E: SUPPLEMENTAL INFORMATION FOR CHAPTER 4

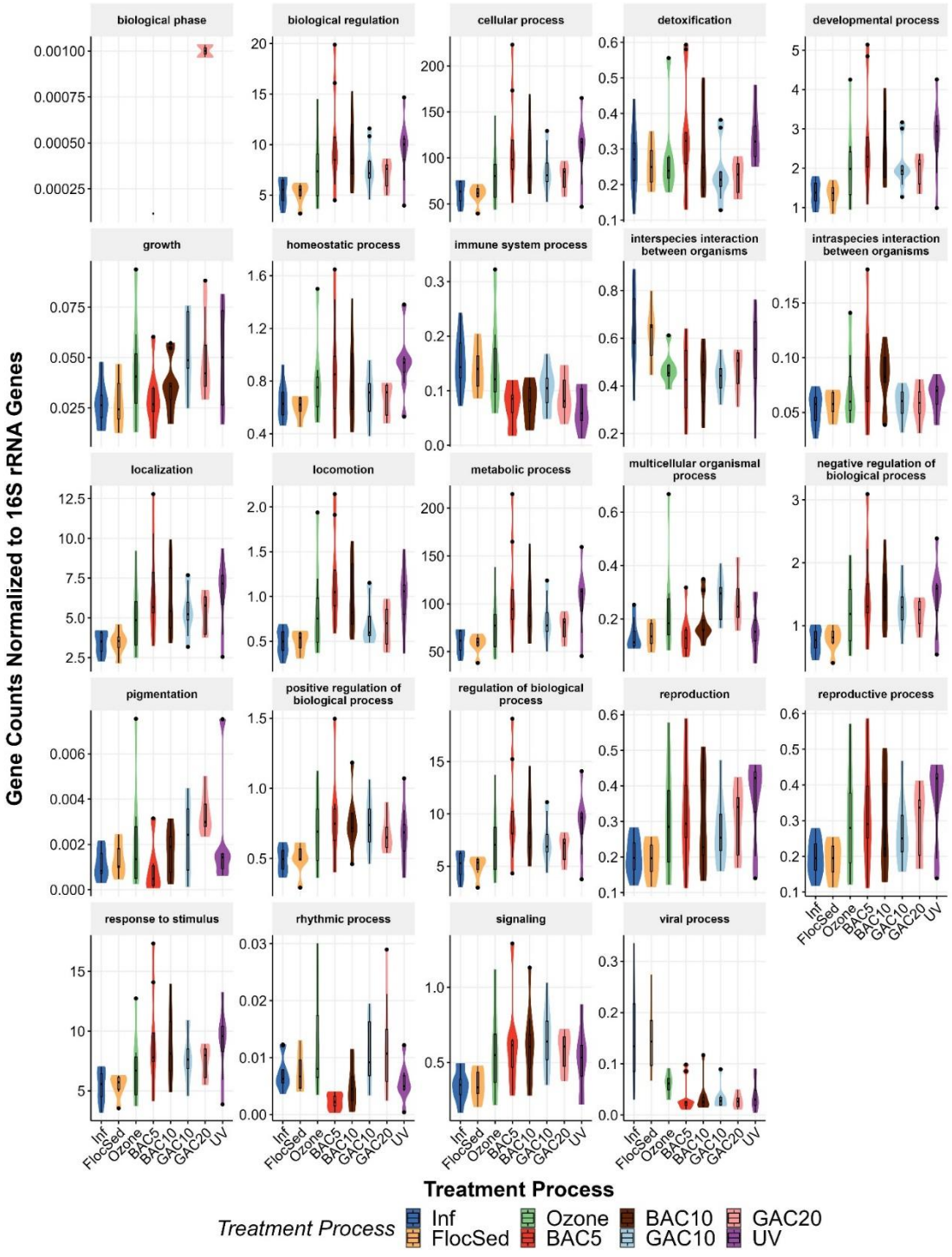
SI Table E.1: Overview of Collected Samples, Operational Conditions, and Upset Events

Sampling Event	Date Sampled or Started	Specific Operational Conditions and Changes	System Upsets and Anomalies	Submitted For Metagenomic Sequencing
-	Before 09/01/2016	**Chlorinated Backwashed Start**	-	
-	09/01/2016	**Monochloramine Dosing Start**	-	
SS-Startup	09/03/2016	-	Limited Replication	
SS-01	11/12/2016	-	-	
SS-02	12/14/2016	-	-	
SS-03	01/19/2017	-	-	
SS-04	02/17/2017	-	-	
-	02/22/2017	**Chlorine Quenching Started (Bisulfite)**	-	
-	03/20/2017	**Chlorinated Backwash Stopped**	-	
SS-05	3/22/2017	-	-	
-	04/01/2017	-	Media Loss BAC5, supplemented with media from BAC10, flow rate adjusted to maintain EBCT	
SUMediaLoss-06	04/12/2017	-	See above	
SS-07	05/17/2017	-	-	
SS-08	06/14/2017	-	-	
SS-09	07/20/2017	-	-	
SS-10	08/23/2017	-	-	
-	09/18/2017	**Nutrient Additions Start** Ammonia (NH4)2SO4 0.5 mg/L as N, Phosphorous H3PO4 0.1 mg/L as P	-	
SSNutrients-11	10/10/2017	1:1 O3:TOC, Thiosulfate Quench, Additional Nutrients (N, P)	-	
-	10/10/2017	**Chiller Start (15° C)**	-	
SSChiller-12	11/14/2017	1:1 O3:TOC, Chilled (15° C), Thiosulfate	-	

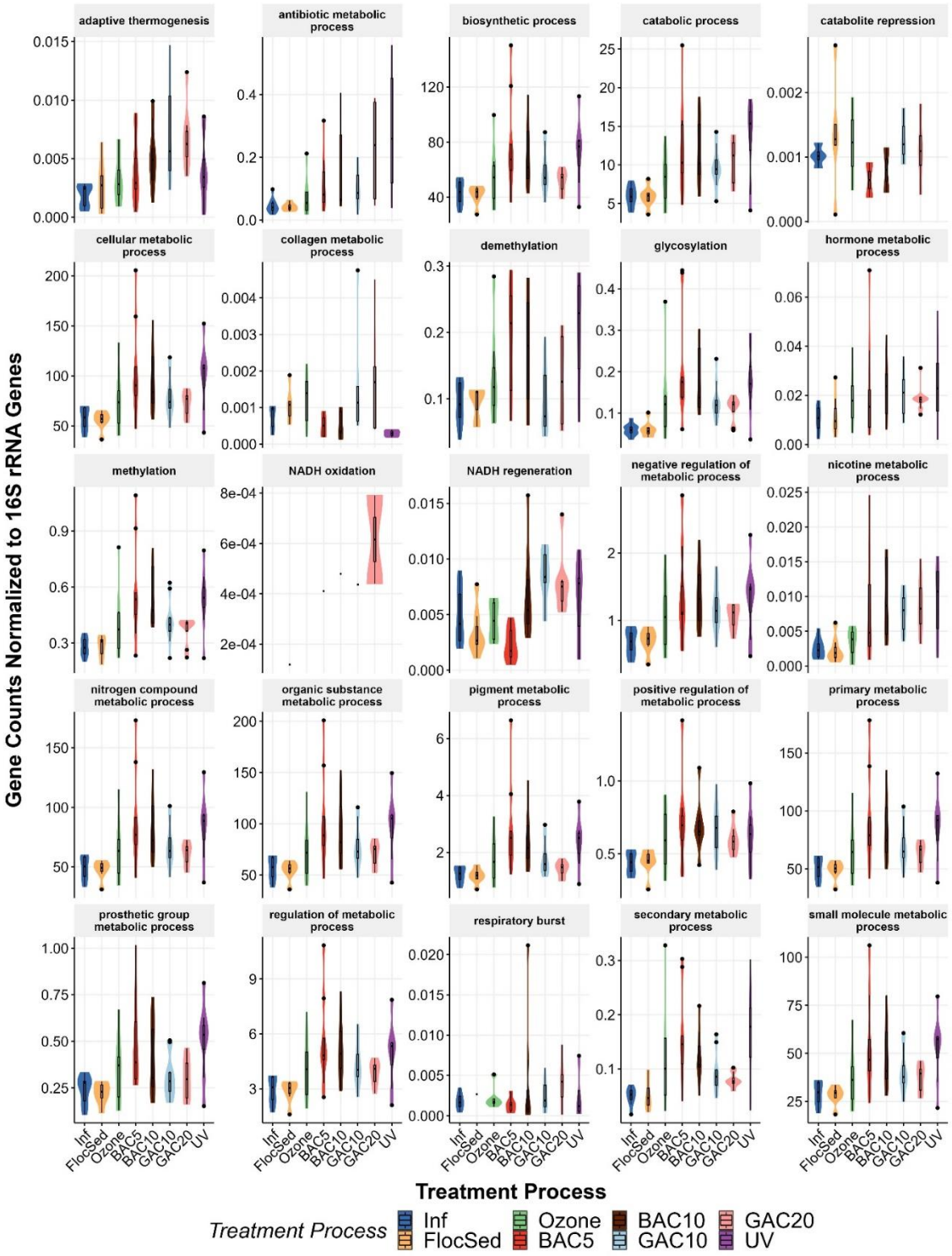
		Quench, Additional Nutrients (N, P)	
SSMediumOzone-13	12/12/2017	0.5:1 O3:TOC, Chilled (15° C), Thiosulfate Quench, Additional Nutrients (N, P)	-
SSLowOzone-14	01/31/2018	0.25:1 O3:TOC, Chilled (15° C), Thiosulfate Quench, Additional Nutrients (N, P)	-
-	02/02/2018	**Return to Bisulfite Quenching**	-
-	02/02/2018	**Filter Aid Polymer Added**	-
SUHighTurbidity-15	02/14/2018	1:1 O3:TOC, Chilled (15° C), Bisulfite Quench, Additional Nutrients (N, P), Filter Polymer Added	Low coagulant dosing (stock issues) lead to large issues with Floc Sed performance. High turbidity event.
-	03/6/2018	**Chlorinated Backwashes Start**	-
SSHighTurbidity-16	03/19/2018	0.8:1 O3:TOC, Chilled (15° C), Bisulfite Quench, Additional Nutrients (N, P), Filter Polymer Added, Chlorine Residual in Backwash	Pilot seeing higher than normal turbidity with elevated nitrite (Ozone consumption)
-	-	**Residual Chlorine Loading Start**	-
SSNoQuench-17	05/19/2018	0.8:1 O3:TOC, Chilled (15° C), No Quench, Additional Nutrients (N, P), Filter Polymer Added	-



SI Figure E.1: NMDS analysis of all annotated functional genes related to metabolic processes not associated with Eukaryotic organisms, based on gene counts normalized to 16S rRNA gene copies. Metabolic functional genes are selected for (A) all treatment processes, and (B) BAC filtration and GAC contracting at tested EBCTs.



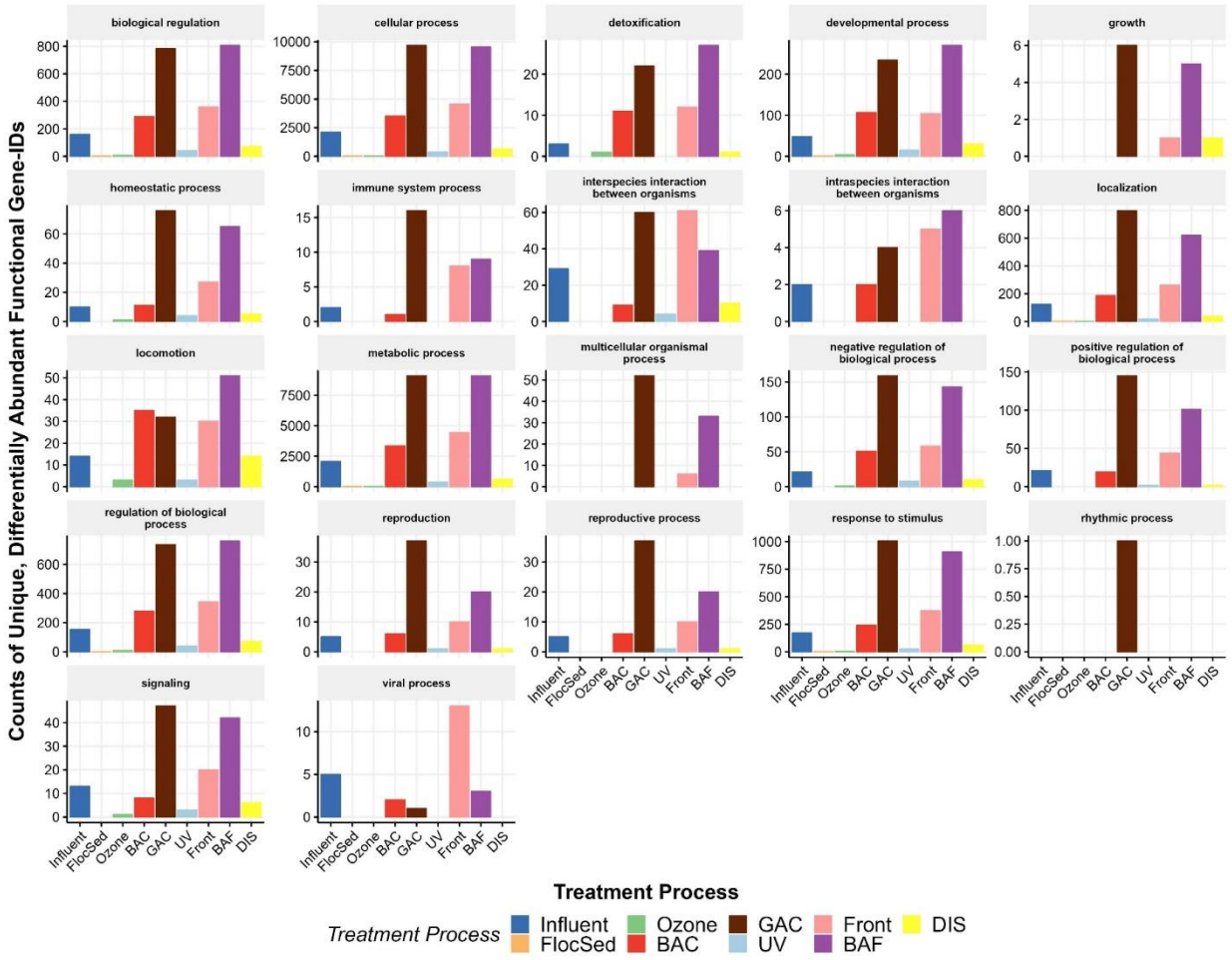
SI Figure E.2: Violin plots of functional genes counts normalized by 16S rRNA genes grouped by treatment process and categories of GO-child terms related to biological process (GO:0008150).



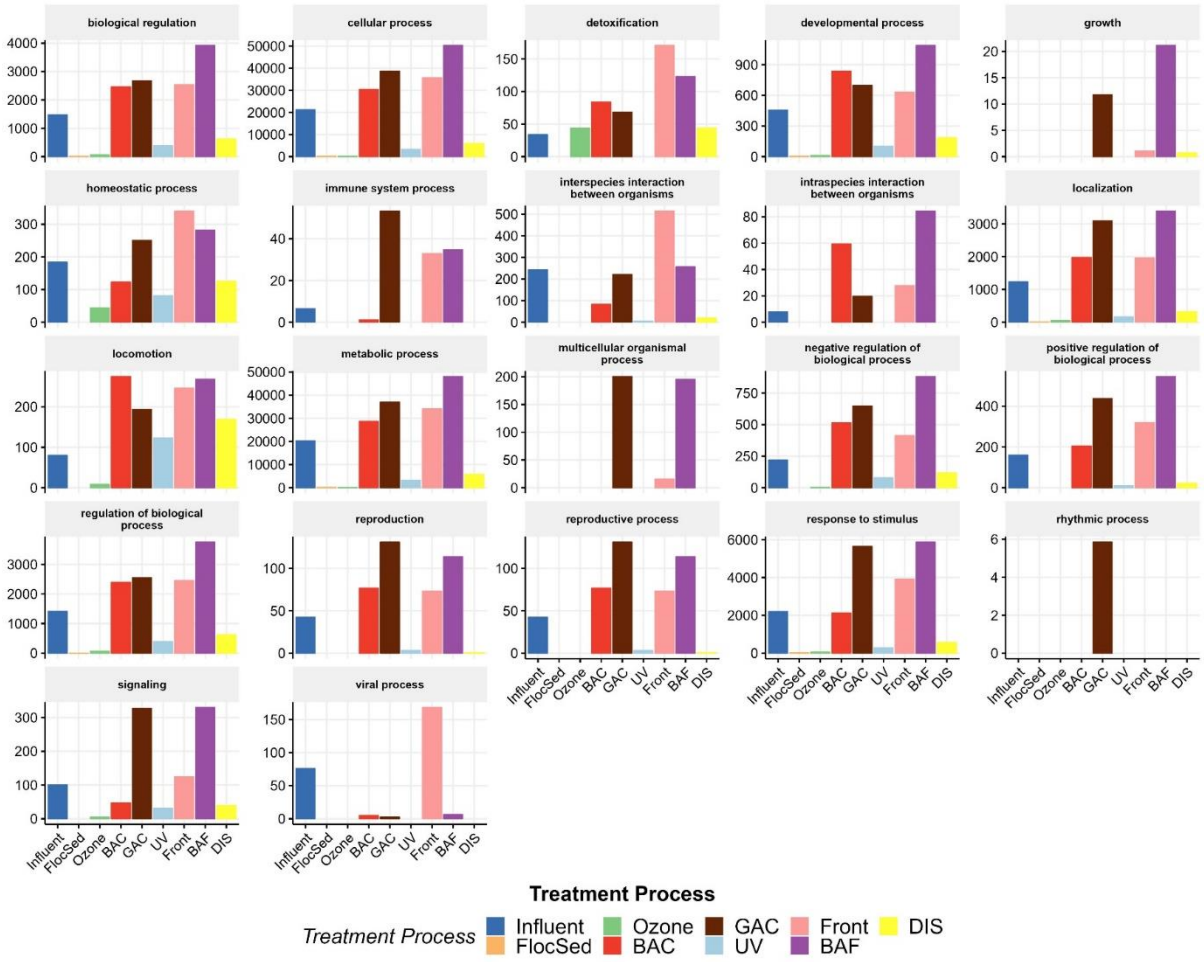
SI Figure E.3: Violin plots of functional genes counts normalized by 16S rRNA genes grouped by treatment process and categories of GO-child terms related to metabolic process (GO:0008152).

SI Table E.2: Summary of genes identified during DESeq2 analysis when comparing individual treatment locations to all other samples.

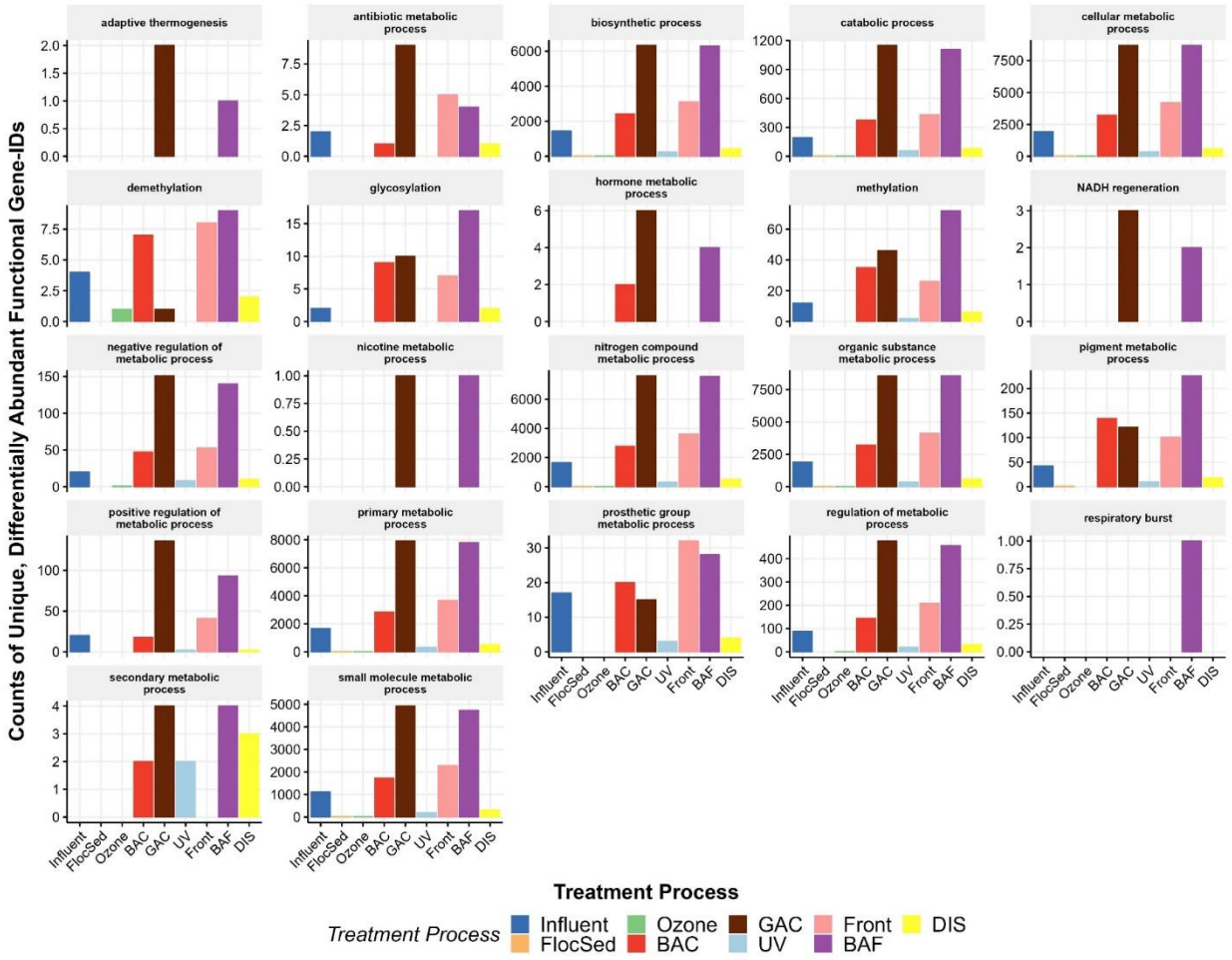
Location	Counts of Unique, Differentially Abudnant Genes				
	Treatment Process	All Other Samples	Total	Percent of Genes at Location	Percent of Genes for all other samples
Influent	2623	8238	10861	24.2%	75.8%
FlocSed	35	1370	1405	2.5%	97.5%
Front	5638	10893	16531	34.1%	65.9%
Ozone	34	19	53	64.2%	35.8%
BAC	4239	8640	12879	32.9%	67.1%
GAC	11698	5182	16880	69.3%	30.7%
BAF	11614	7099	18713	62.1%	37.9%
UV	464	1092	1556	29.8%	70.2%
DIS	794	2344	3138	25.3%	74.7%



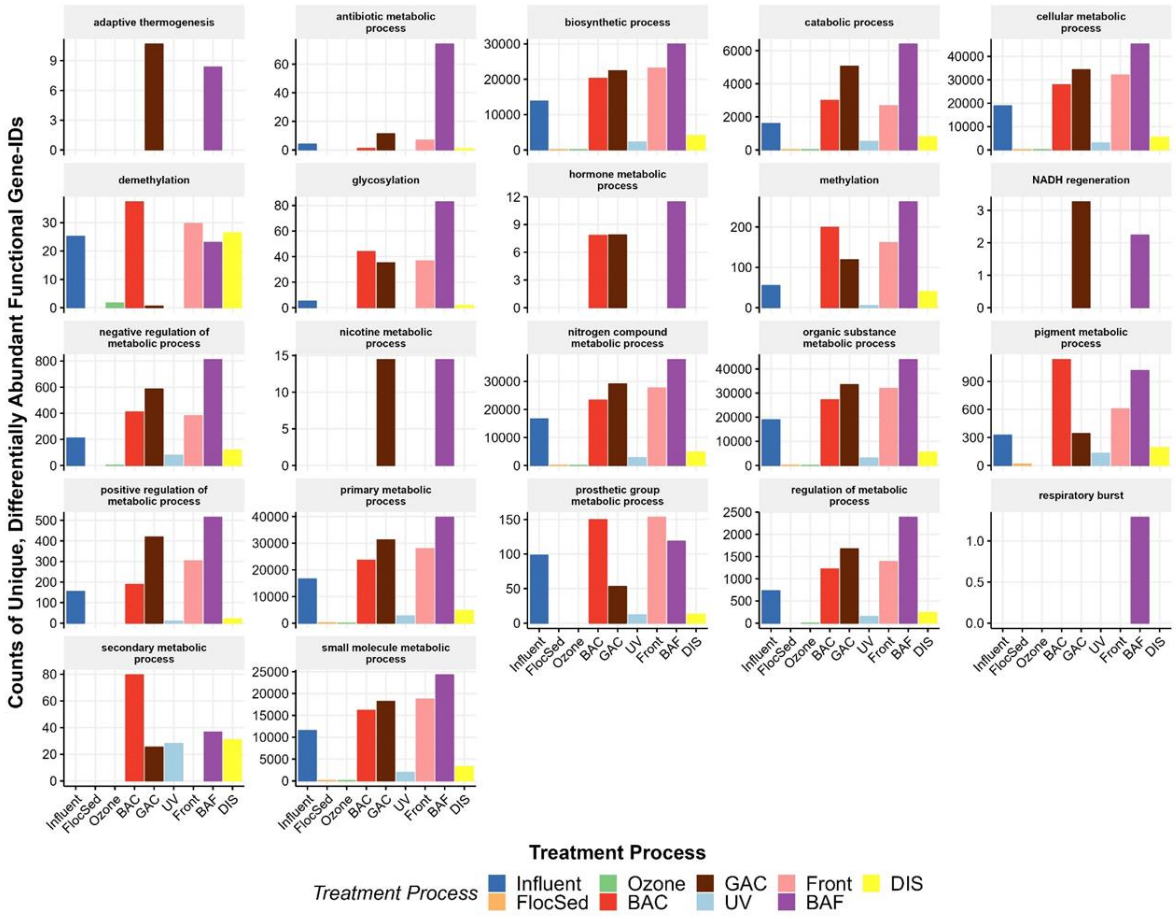
SI Figure E.4: Counts of unique, differentially abundant genes determined by DESeq2 with an adjusted p-value cutoff of 0.05. Gene counts are grouped by treatment process and categories of GO-child terms related to biological process (GO:0008150). Differential genes were determined by comparisons between individual treatment process and all other samples.



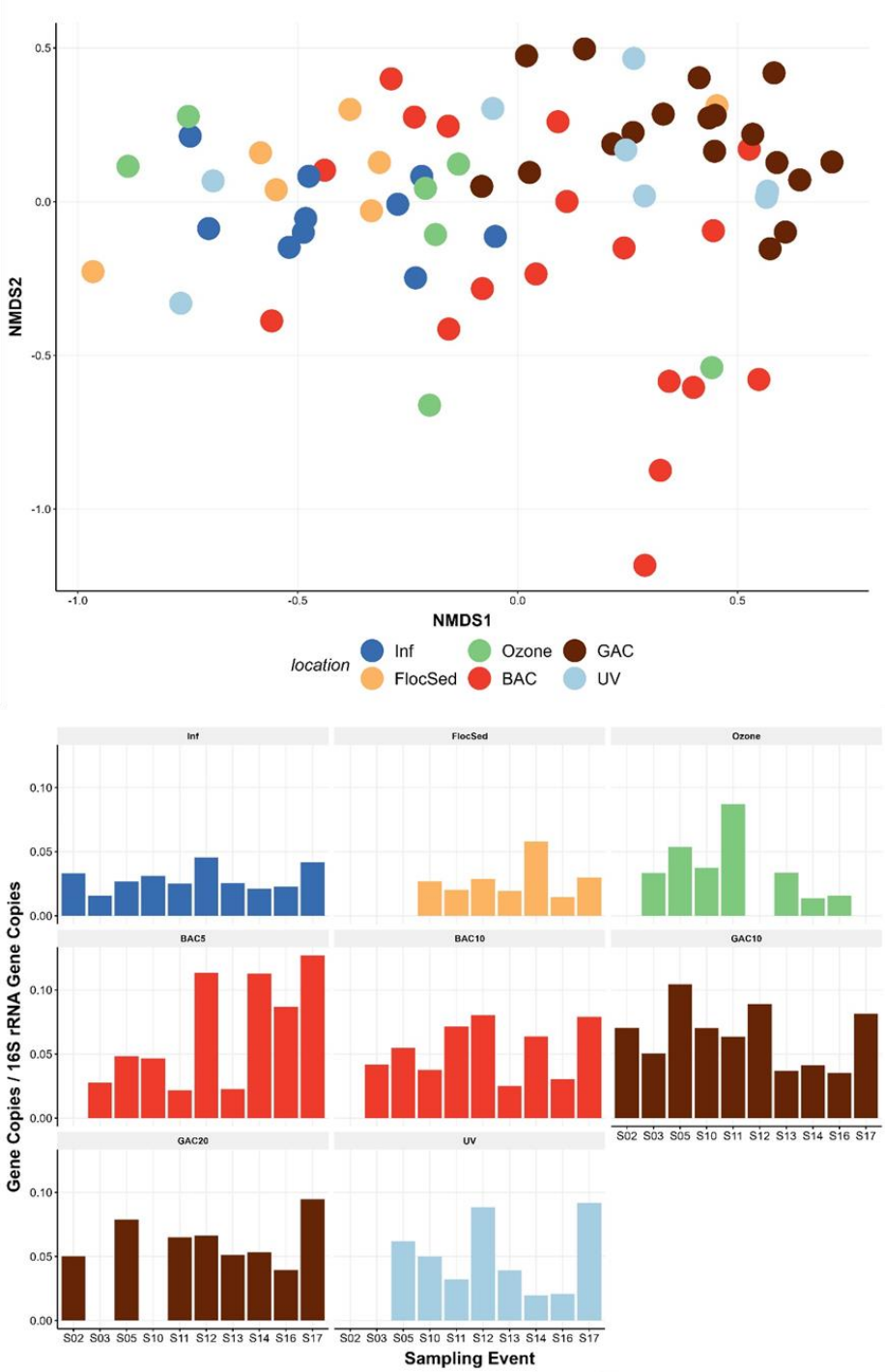
SI Figure E.5: Mean abundances of unique, differentially abundant genes determined by DESeq2 with an adjusted p-value cutoff of 0.05. Gene counts are grouped by treatment process and categories of GO-child terms related to biological process (GO:0008150). Differential genes were determined by comparisons between individual treatment process and all other samples.



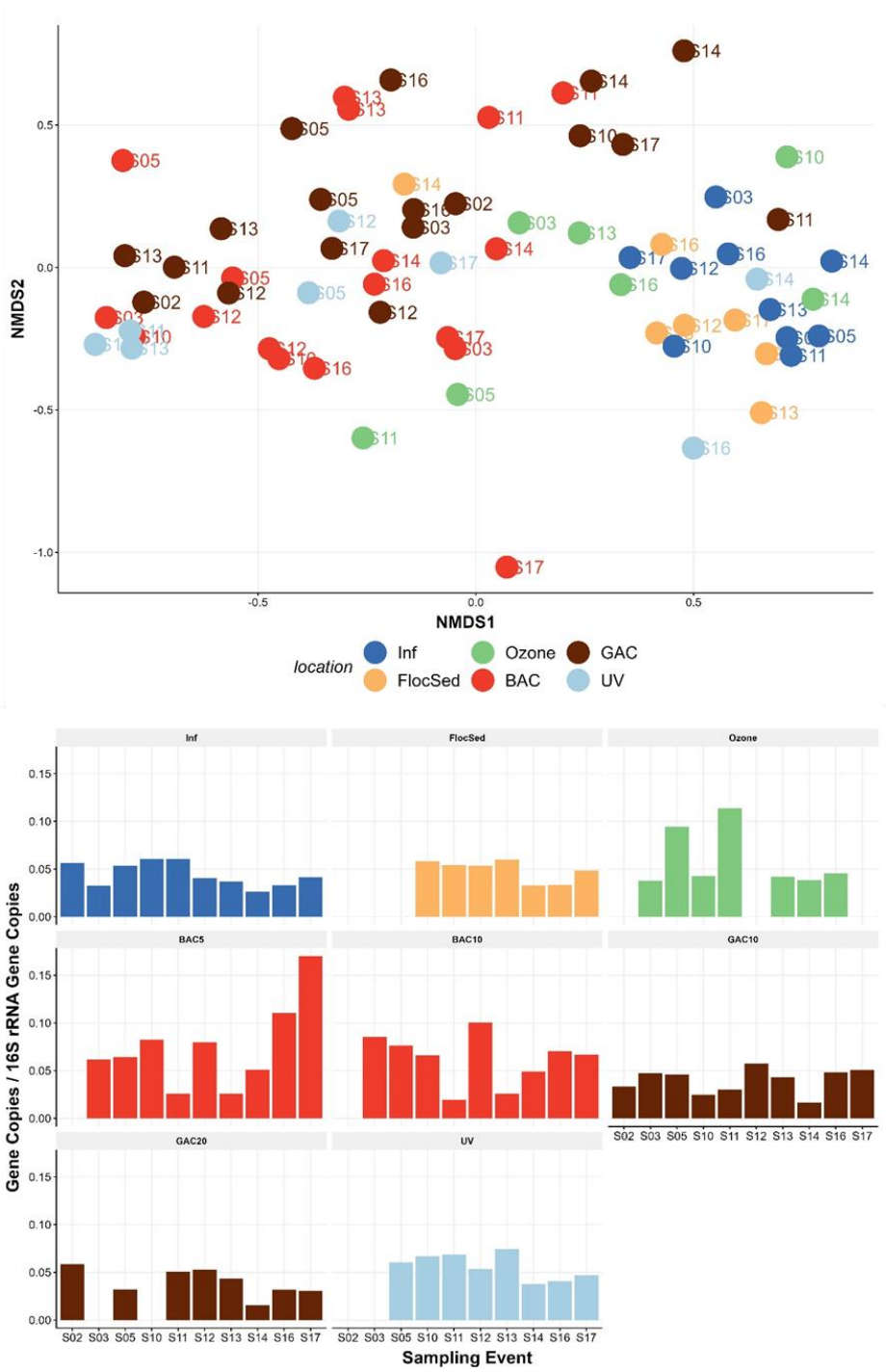
SI Figure E.6: Counts of unique, differentially abundant genes determined by DESeq2 with an adjusted p-value cutoff of 0.05. Gene counts are grouped by treatment process and categories of GO-child terms related to metabolic process (GO:0008152). Differential genes were determined by comparisons between individual treatment process and all other samples.



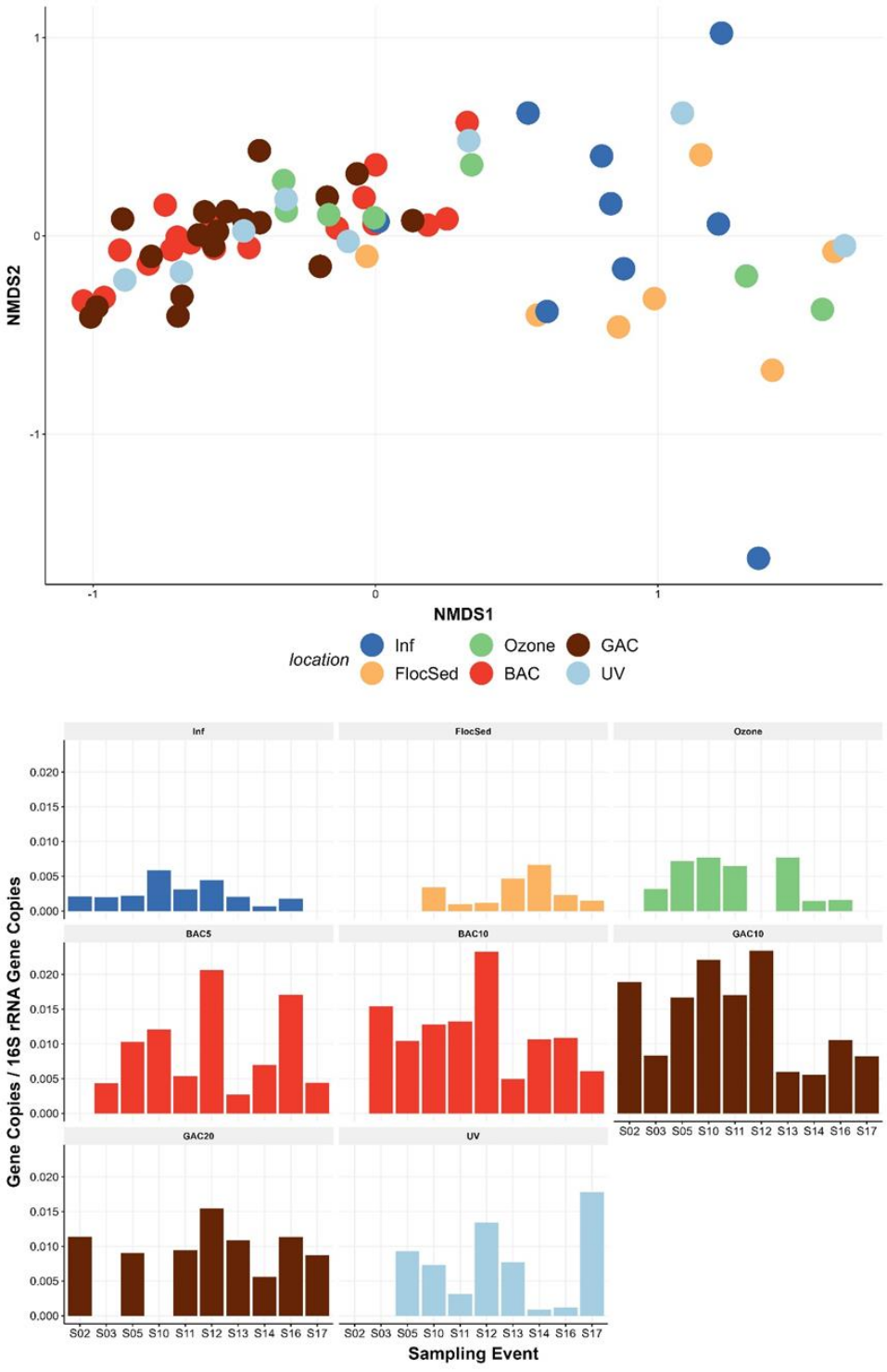
SI Figure E.7: Mean abundances of unique, differentially abundant genes determined by DESeq2 with an adjusted p-value cutoff of 0.05. Gene counts are grouped by treatment process and categories of GO-child terms related to metabolic process (GO:0008152). Differential genes were determined by comparisons between individual treatment process and all other samples.



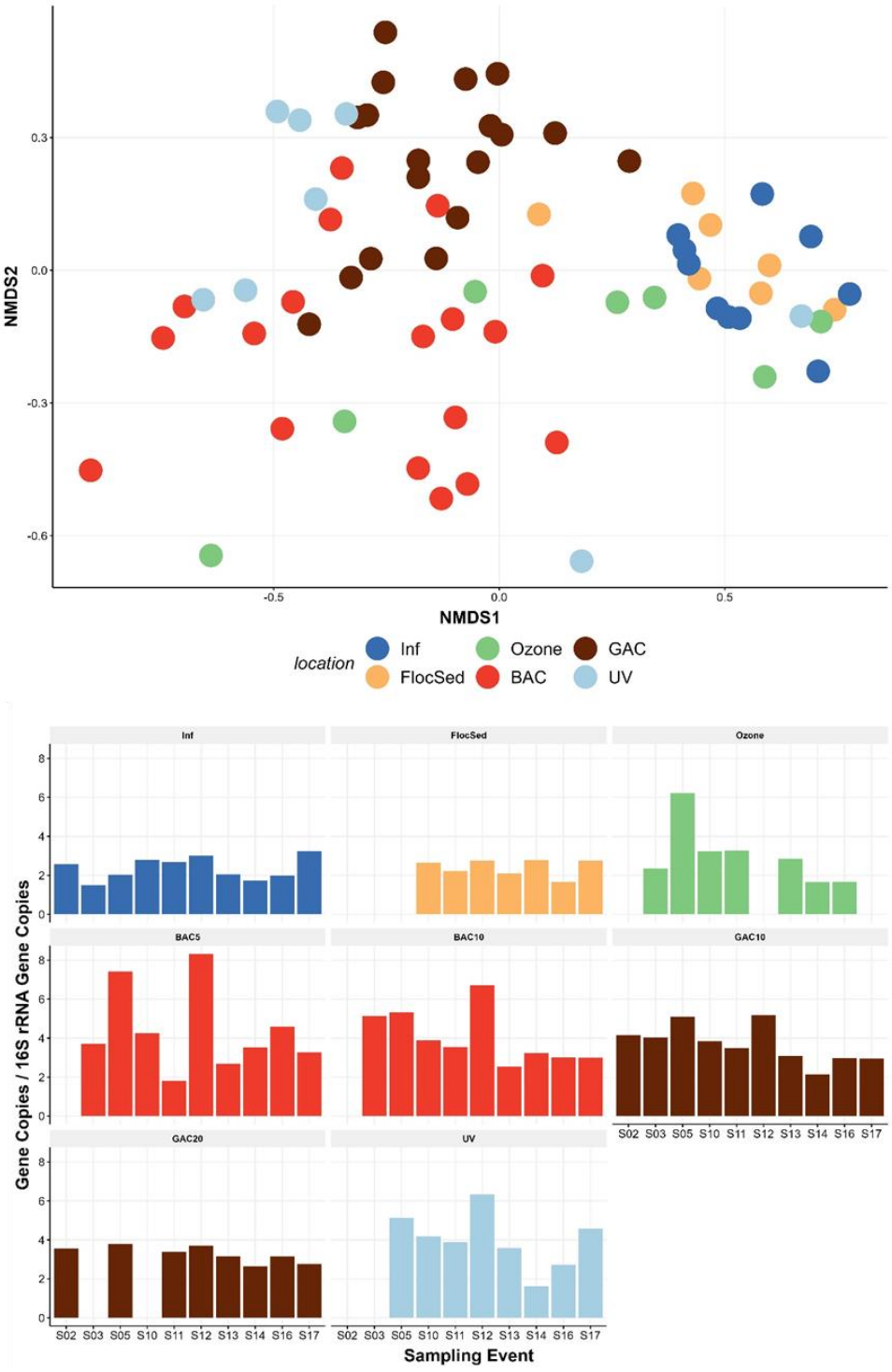
SI Figure E.8: Supplemental figures associated with Horizontal Gene Transfer's gene counts normalized by 16S rRNA gene. Includes (A) Bray Curtis NDMS plot and (B) bar plot of categorized relative abundances grouped by sampling events.



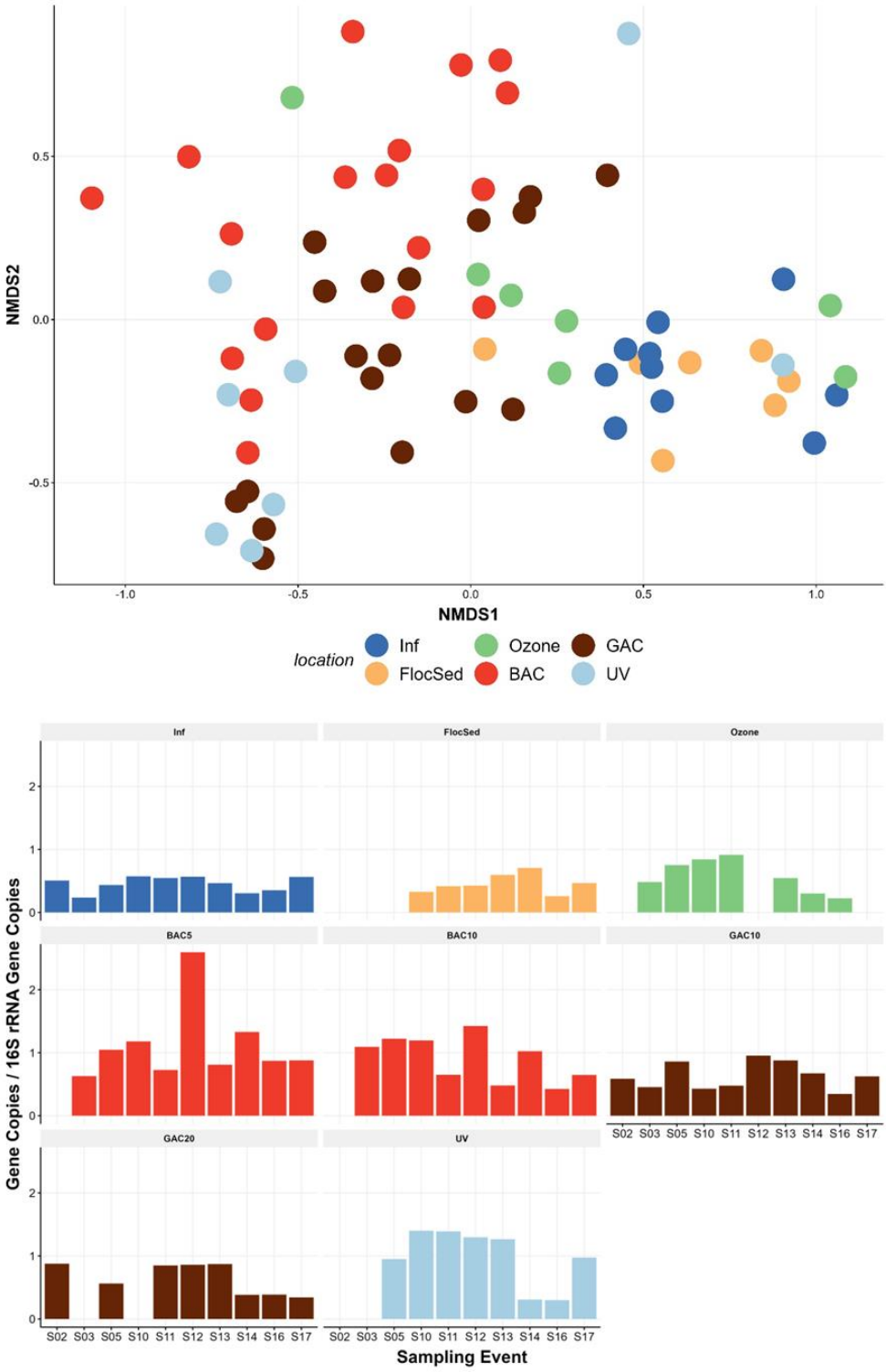
SI Figure E.9: Supplemental figures associated with Biofilm Formation's gene counts normalized by 16S rRNA gene. Includes (A) Bray Curtis NDMS plot and (B) bar plot of categorized relative abundances grouped by sampling events.



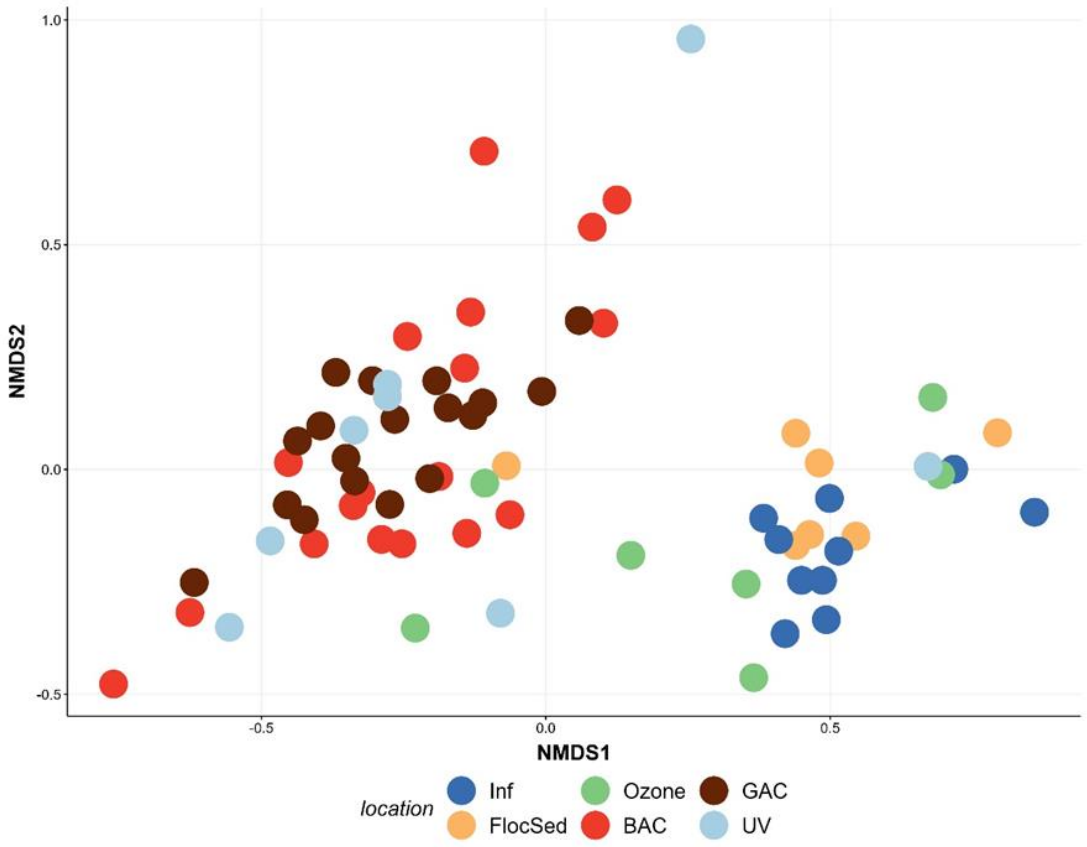
SI Figure E.10: Supplemental figures associated with Response to Ozone's gene counts normalized by 16S rRNA gene. Includes (A) Bray Curtis NDMS plot and (B) bar plot of categorized relative abundances grouped by sampling events.



SI Figure E.11: Supplemental figures associated with DNA Repair's gene counts normalized by 16S rRNA gene. Includes (A) Bray Curtis NDMS plot and (B) bar plot of categorized relative abundances grouped by sampling events.



SI Figure E.12: Supplemental figures associated with Monooxygenase activity's gene counts normalized by 16S rRNA gene. Includes (A) Bray Curtis NDMS plot and (B) bar plot of categorized relative abundances grouped by sampling events.



SI Figure E.13: Supplemental NDMS plot associated with Nitrogen Cycling gene counts normalized by 16S rRNA gene.

SI Table E.3: Average relative abundances of the phylum Cyanobacteria at specific treatment locations determined via 16S rRNA gene amplicon sequencing.

Treatment Location	Average Relative Abundance of Cyanobacteria
Influent	0.16%
Ozone	0.47%
BAC	0.26%
GAC	0.64%

APPENDIX F: SUPPLEMENTAL INFORMATION FOR CHAPTER 5

The following paragraph outlines the selection of water quality parameters and water quality data. Because water quality data were collected periodically, sampling dates were used to sort and identify data collected around each sampling event, specifically the day of and the three days preceding and following each sampling date. Because all parameters were not measured every day of operation, selection criteria were developed that prioritized: 1) data collected on the day of sampling, 2) average of the day before and after sampling, 3) the only value in the data range if only one exists, or 4) a blank value. After selection, the data were used to benchmark BDOC results to water quality characteristics. Operational data, the data and parameters collected during the continued operation of the treatment skids from online analyzers, was collected in approximately 5-minute intervals. Data exported from the treatment skids were utilized to calculate daily summary statistics including averages, standard deviations, minima, and maxima, which were used to benchmark treatment operational to subsequent results.

SI Table F.1: Percent of sequences classified at various levels of classification.

Percent of Classified Reads					
Sample Location	Phylum	Class	Order	Family	Genus
Influent	99.5%	98.7%	93.6%	86.1%	23.3%
FlocSed	99.4%	98.5%	91.0%	84.2%	24.4%
Ozone	99.4%	99.0%	95.7%	87.8%	40.0%
BAC5	98.7%	98.2%	94.2%	87.9%	54.1%
BAC10	99.2%	98.4%	91.7%	80.6%	53.2%
GAC10	99.3%	97.6%	79.6%	58.8%	36.0%
GAC20	98.9%	96.9%	75.6%	54.7%	30.1%
UV	99.4%	98.2%	90.7%	80.4%	36.7%
Average	99.2%	98.2%	89.0%	77.6%	37.2%

SI Table F.2: Percent of sequences unclassified at various levels of classification.

Percentage of Unclassified Reads					
Sample Location	Phylum	Class	Order	Family	Genus
Influent	0.5%	1.3%	6.4%	13.9%	76.7%
FlocSed	0.6%	1.5%	9.0%	15.8%	75.6%
Ozone	0.6%	1.0%	4.3%	12.2%	60.0%
BAC5	1.3%	1.8%	5.8%	12.1%	45.9%
BAC10	0.8%	1.6%	8.3%	19.4%	46.8%
GAC10	0.7%	2.4%	20.4%	41.2%	64.0%
GAC20	1.1%	3.1%	24.4%	45.3%	69.9%
UV	0.6%	1.8%	9.3%	19.6%	63.3%
Average	0.8%	1.8%	11.0%	22.4%	62.8%

SI Table F.3: Full Pilot Anosim Results

Anosim_weighted_3500	signif	permutations	statistic
Anosim.BAC5.YR_Ammonia_Additions	0.000999001	1000	0.3375
Anosim.BAC5.YR_Polymer_Addition	0.037962038	1000	0.1878
Anosim.BAC5.YR_Quenching_Type	0.000999001	1000	0.3151
Anosim.BAC5.YR_SampleEvent	0.000999001	1000	1.0000
Anosim_Unweighted_3500	signif	permutations	statistic
Anosim.BAC5.YR_Ammonia_Additions	0.000999001	1000	0.3195
Anosim.BAC5.YR_Polymer_Addition	0.00999001	1000	0.2486
Anosim.BAC5.YR_Quenching_Type	0.001998002	1000	0.2259
Anosim.BAC5.YR_SampleEvent	0.000999001	1000	0.9621

SI Table F.4: Full Pilot Adonis Results

Adonis_Weighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.BAC5.YR_BackwashProximity_HN	1	0.3477	0.3477	7.0702	0.1261	0.0003
Adonis.BAC5.YR_OzoneDose_Filter	1	0.2685	0.2685	5.2862	0.0974	0.0001
Adonis.BAC5.YR_OzoneTOC_Adjusted_Filter	1	0.3056	0.3056	6.1078	0.1108	0.0001
Adonis.BAC5.YR_OzoneTOC_Filter	1	0.3835	0.3835	7.9161	0.1391	0.0001
Adonis.BAC5.YR_Temperature	1	0.3627	0.3627	7.4214	0.1315	0.0001
Adonis.BAC5.YR_Time_Days	1	0.3170	0.3170	6.3649	0.1150	0.0001
Adonis_Unweighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.BAC5.YR_BackwashProximity_HN	1	0.7223	0.7223	2.9321	0.0565	0.0016
Adonis.BAC5.YR_OzoneDose_Filter	1	1.6008	1.6008	7.0081	0.1251	0.0001
Adonis.BAC5.YR_OzoneTOC_Adjusted_Filter	1	1.5262	1.5262	6.6375	0.1193	0.0001
Adonis.BAC5.YR_OzoneTOC_Filter	1	1.9068	1.9068	8.5824	0.1490	0.0001
Adonis.BAC5.YR_Temperature	1	1.8795	1.8795	8.4385	0.1469	0.0001
Adonis.BAC5.YR_Time_Days	1	0.9111	0.9111	3.7574	0.0712	0.0003

SI Table F.5: Influent and FlocSed Anosim Results

Anosim_weighted_3500	signif	permutations	statistic
Anosim.IFS.YR_Location_adjusted	0.006993007	1000	0.0897
Anosim.IFS.YR_SampleEvent	0.000999001	1000	0.8375
Anosim_Unweighted_3500	signif	permutations	statistic
Anosim.IFS.YR_Location_adjusted	0.000999001	1000	0.1890
Anosim.IFS.YR_SampleEvent	0.000999001	1000	0.9339

SI Table F.6: Influent and FlocSed Adonis Results

Adonis_Weighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.IFS.YR_Temperature	1	0.1007	0.1007	19.0764	0.1815	0.0001
Adonis.IFS.YR_Time_Days	1	0.0319	0.0319	5.2490	0.0575	0.0006
Adonis_Unweighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.IFS.YR_Temperature	1	2.3028	2.3028	12.5059	0.1270	0.0001
Adonis.IFS.YR_Time_Days	1	1.0346	1.0346	5.2018	0.0570	0.0001

SI Table F.7: Ozone Anosim Results

Anosim_weighted_3500	signif	permutations	statistic
Anosim.Ozone.YR_SampleEvent	0.000999001	1000	0.9846
Anosim_Unweighted_3500	signif	permutations	statistic
Anosim.Ozone.YR_SampleEvent	0.000999001	1000	0.9899

SI Table F.8: Ozone Adonis Results

Adonis_Weighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.Ozone.YR_OzoneDose_Ozone	1	0.3855	0.3855	9.4595	0.1675	0.0001
Adonis.Ozone.YR_OzoneTOC_Adjusted_Ozone	1	0.5320	0.5320	14.1356	0.2312	0.0001
Adonis.Ozone.YR_OzoneTOC_Ozone	1	0.4744	0.4744	12.2072	0.2062	0.0001
Adonis.Ozone.YR_Temperature	1	0.5720	0.5720	15.5515	0.2486	0.0001
Adonis.Ozone.YR_Time_Days	1	0.1546	0.1546	3.3848	0.0672	0.0173
Adonis_Unweighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.Ozone.YR_OzoneDose_Ozone	1	1.3901	1.3901	6.2038	0.1166	0.0001

Adonis.Ozone.YR_OzoneTOC_Adjusted_Ozone	1	1.6398	1.6398	7.4957	0.1375	0.0001
Adonis.Ozone.YR_OzoneTOC_Ozone	1	1.6474	1.6474	7.5360	0.1382	0.0001
Adonis.Ozone.YR_Temperature	1	1.6635	1.6635	7.6216	0.1395	0.0001
Adonis.Ozone.YR_Time_Days	1	0.6053	0.6053	2.5138	0.0508	0.0012

SI Table F.9: BAC5 Anosim Results

Anosim_weighted_3500	signif	permutations	statistic
Anosim.BAC5.YR_Ammonia_Additions	0.000999001	1000	0.3375
Anosim.BAC5.YR_Polymer_Addition	0.037962038	1000	0.1878
Anosim.BAC5.YR_Quenching_Type	0.000999001	1000	0.3151
Anosim.BAC5.YR_SampleEvent	0.000999001	1000	1.0000
Anosim_Unweighted_3500	signif	permutations	statistic
Anosim.BAC5.YR_Ammonia_Additions	0.000999001	1000	0.3195
Anosim.BAC5.YR_Polymer_Addition	0.00999001	1000	0.2486
Anosim.BAC5.YR_Quenching_Type	0.001998002	1000	0.2259
Anosim.BAC5.YR_SampleEvent	0.000999001	1000	0.9621

SI Table F.10: BAC5 Adonis Results

Adonis_Weighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.BAC5.YR_BackwashProximity_HN	1	0.3477	0.3477	7.0702	0.1261	0.0003
Adonis.BAC5.YR_OzoneDose_Filter	1	0.2685	0.2685	5.2862	0.0974	0.0001
Adonis.BAC5.YR_OzoneTOC_Adjusted_Filter	1	0.3056	0.3056	6.1078	0.1108	0.0001
Adonis.BAC5.YR_OzoneTOC_Filter	1	0.3835	0.3835	7.9161	0.1391	0.0001
Adonis.BAC5.YR_Temperature	1	0.3627	0.3627	7.4214	0.1315	0.0001
Adonis.BAC5.YR_Time_Days	1	0.3170	0.3170	6.3649	0.1150	0.0001
Adonis_Unweighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.BAC5.YR_BackwashProximity_HN	1	0.7223	0.7223	2.9321	0.0565	0.0016
Adonis.BAC5.YR_OzoneDose_Filter	1	1.6008	1.6008	7.0081	0.1251	0.0001

Adonis.BAC5.YR_OzoneTOC_Adjusted_Filter	1	1.5262	1.5262	6.6375	0.1193	0.0001
Adonis.BAC5.YR_OzoneTOC_Filter	1	1.9068	1.9068	8.5824	0.1490	0.0001
Adonis.BAC5.YR_Temperature	1	1.8795	1.8795	8.4385	0.1469	0.0001
Adonis.BAC5.YR_Time_Days	1	0.9111	0.9111	3.7574	0.0712	0.0003

SI Table F.11: BAC10 Anosim Results

Anosim_weighted_3500	signif	permutations	statistic
Anosim.BAC10.YR_Ammonia_Additions	0.000999001	1000	0.4764
Anosim.BAC10.YR_Polymer_Addition	0.000999001	1000	0.4925
Anosim.BAC10.YR_Quenching_Type	0.004995005	1000	0.2318
Anosim.BAC10.YR_SampleEvent	0.000999001	1000	0.9900
Anosim_Unweighted_3500	signif	permutations	statistic
Anosim.BAC10.YR_Ammonia_Additions	0.000999001	1000	0.5387
Anosim.BAC10.YR_Polymer_Addition	0.007992008	1000	0.2871
Anosim.BAC10.YR_Quenching_Type	0.000999001	1000	0.3547
Anosim.BAC10.YR_SampleEvent	0.000999001	1000	0.9469

SI Table F.12: BAC10 Adonis Results

Adonis_Weighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.BAC10.YR_BackwashProximity_HN	1	0.0652	0.0652	1.5822	0.0313	0.1547
Adonis.BAC10.YR_OzoneDose_Filter	1	0.4243	0.4243	12.5274	0.2036	0.0001
Adonis.BAC10.YR_OzoneTOC_Adjusted_Filter	1	0.5368	0.5368	17.0010	0.2576	0.0001
Adonis.BAC10.YR_OzoneTOC_Filter	1	0.4471	0.4471	13.3840	0.2145	0.0001
Adonis.BAC10.YR_Temperature	1	0.4530	0.4530	13.6075	0.2173	0.0001
Adonis.BAC10.YR_Time_Days	1	0.3662	0.3662	10.4457	0.1757	0.0001
Adonis_Unweighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.BAC10.YR_BackwashProximity_HN	1	0.6476	0.6476	3.0717	0.0590	0.0015
Adonis.BAC10.YR_OzoneDose_Filter	1	1.5038	1.5038	7.7780	0.1370	0.0001

Adonis.BAC10.YR_OzoneTOC_Adjusted_Filter	1	1.5933	1.5933	8.3199	0.1451	0.0001
Adonis.BAC10.YR_OzoneTOC_Filter	1	1.6664	1.6664	8.7699	0.1518	0.0001
Adonis.BAC10.YR_Temperature	1	1.8438	1.8438	9.8920	0.1680	0.0001
Adonis.BAC10.YR_Time_Days	1	1.4470	1.4470	7.4398	0.1318	0.0001

SI Table F.13: GAC10 Anosim Results

Anosim_weighted_3500	signif	permutations	statistic
Anosim.GAC10.YR_Ammonia_Additions	0.000999001	1000	0.4413
Anosim.GAC10.YR_Polymer_Addition	0.000999001	1000	0.5074
Anosim.GAC10.YR_Quenching_Type	0.000999001	1000	0.2557
Anosim.GAC10.YR_SampleEvent	0.000999001	1000	0.9999
Anosim_Unweighted_3500	signif	permutations	statistic
Anosim.GAC10.YR_Ammonia_Additions	0.000999001	1000	0.5422
Anosim.GAC10.YR_Polymer_Addition	0.000999001	1000	0.5309
Anosim.GAC10.YR_Quenching_Type	0.000999001	1000	0.2579
Anosim.GAC10.YR_SampleEvent	0.000999001	1000	0.9778

SI Table F.14: GAC10 Adonis Results

Adonis_Weighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.GAC10.YR_BackwashProximity_Skid	1	0.1916	0.1916	4.0908	0.0771	0.0009
Adonis.GAC10.YR_OzoneDose_Filter	1	0.3651	0.3651	8.4345	0.1469	0.0001
Adonis.GAC10.YR_OzoneTOC_Adjusted_Filter	1	0.3729	0.3729	8.6446	0.1500	0.0001
Adonis.GAC10.YR_OzoneTOC_Filter	1	0.4077	0.4077	9.6111	0.1640	0.0001
Adonis.GAC10.YR_Temperature	1	0.3985	0.3985	9.3536	0.1603	0.0001
Adonis.GAC10.YR_Time_Days	1	0.3736	0.3736	8.6659	0.1503	0.0001
Adonis_Unweighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.GAC10.YR_BackwashProximity_Skid	1	0.5715	0.5715	3.1572	0.0605	0.0004
Adonis.GAC10.YR_OzoneDose_Filter	1	1.1959	1.1959	7.1072	0.1267	0.0001

Adonis.GAC10.YR_OzoneTOC_Adjusted_Filter	1	1.3186	1.3186	7.9548	0.1397	0.0001
Adonis.GAC10.YR_OzoneTOC_Filter	1	1.3851	1.3851	8.4250	0.1467	0.0001
Adonis.GAC10.YR_Temperature	1	1.3479	1.3479	8.1608	0.1428	0.0001
Adonis.GAC10.YR_Time_Days	1	1.3710	1.3710	8.3244	0.1452	0.0001

SI Table F.15: GAC20 Anosim Results

Anosim_weighted_3500	signif	permutations	statistic
Anosim.GAC20.YR_Ammonia_Additions	0.000999001	1000	0.396362627
Anosim.GAC20.YR_Polymer_Addition	0.000999001	1000	0.739889706
Anosim.GAC20.YR_Quenching_Type	0.066933067	1000	0.099412585
Anosim.GAC20.YR_SampleEvent	0.000999001	1000	0.963255098
Anosim_Unweighted_3500	signif	permutations	statistic
Anosim.GAC20.YR_Ammonia_Additions	0.000999001	1000	0.625503263
Anosim.GAC20.YR_Polymer_Addition	0.000999001	1000	0.56479439
Anosim.GAC20.YR_Quenching_Type	0.000999001	1000	0.29210225
Anosim.GAC20.YR_SampleEvent	0.000999001	1000	0.986148519

SI Table F.16: GAC20 Adonis Results

Adonis_Weighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.GAC20.YR_BackwashProximity_Skid	1	0.2368	0.2368	4.6879	0.0907	0.0005
Adonis.GAC20.YR_OzoneDose_Filter	1	0.2484	0.2484	4.9422	0.0951	0.0001
Adonis.GAC20.YR_OzoneTOC_Adjusted_Filter	1	0.3244	0.3244	6.6694	0.1243	0.0001
Adonis.GAC20.YR_OzoneTOC_Filter	1	0.2850	0.2850	5.7584	0.1091	0.0001
Adonis.GAC20.YR_Temperature	1	0.2854	0.2854	5.7695	0.1093	0.0001
Adonis.GAC20.YR_Time_Days	1	0.3641	0.3641	7.6167	0.1395	0.0001
Adonis_Unweighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.GAC20.YR_BackwashProximity_Skid	1	0.6484	0.6484	3.9216	0.0770	0.0001
Adonis.GAC20.YR_OzoneDose_Filter	1	0.9012	0.9012	5.6340	0.1070	0.0001
Adonis.GAC20.YR_OzoneTOC_Adjusted_Filter	1	1.0884	1.0884	6.9779	0.1293	0.0001
Adonis.GAC20.YR_OzoneTOC_Filter	1	1.0430	1.0430	6.6452	0.1239	0.0001
Adonis.GAC20.YR_Temperature	1	1.0061	1.0061	6.3785	0.1195	0.0001
Adonis.GAC20.YR_Time_Days	1	1.3841	1.3841	9.2463	0.1644	0.0001

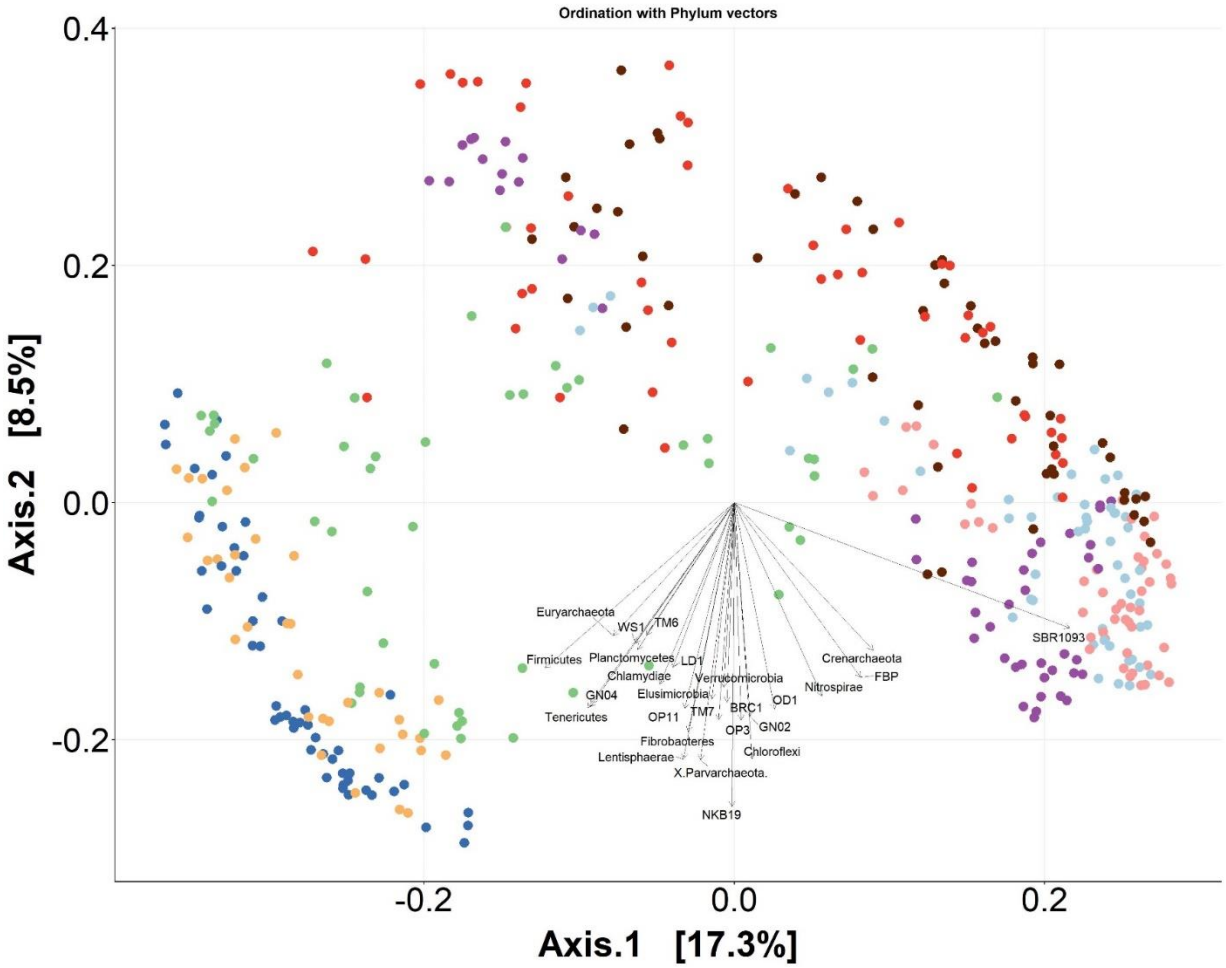
SI Table F.17: UV Anosim Results

Anosim_weighted_3500	signif	permutations	statistic

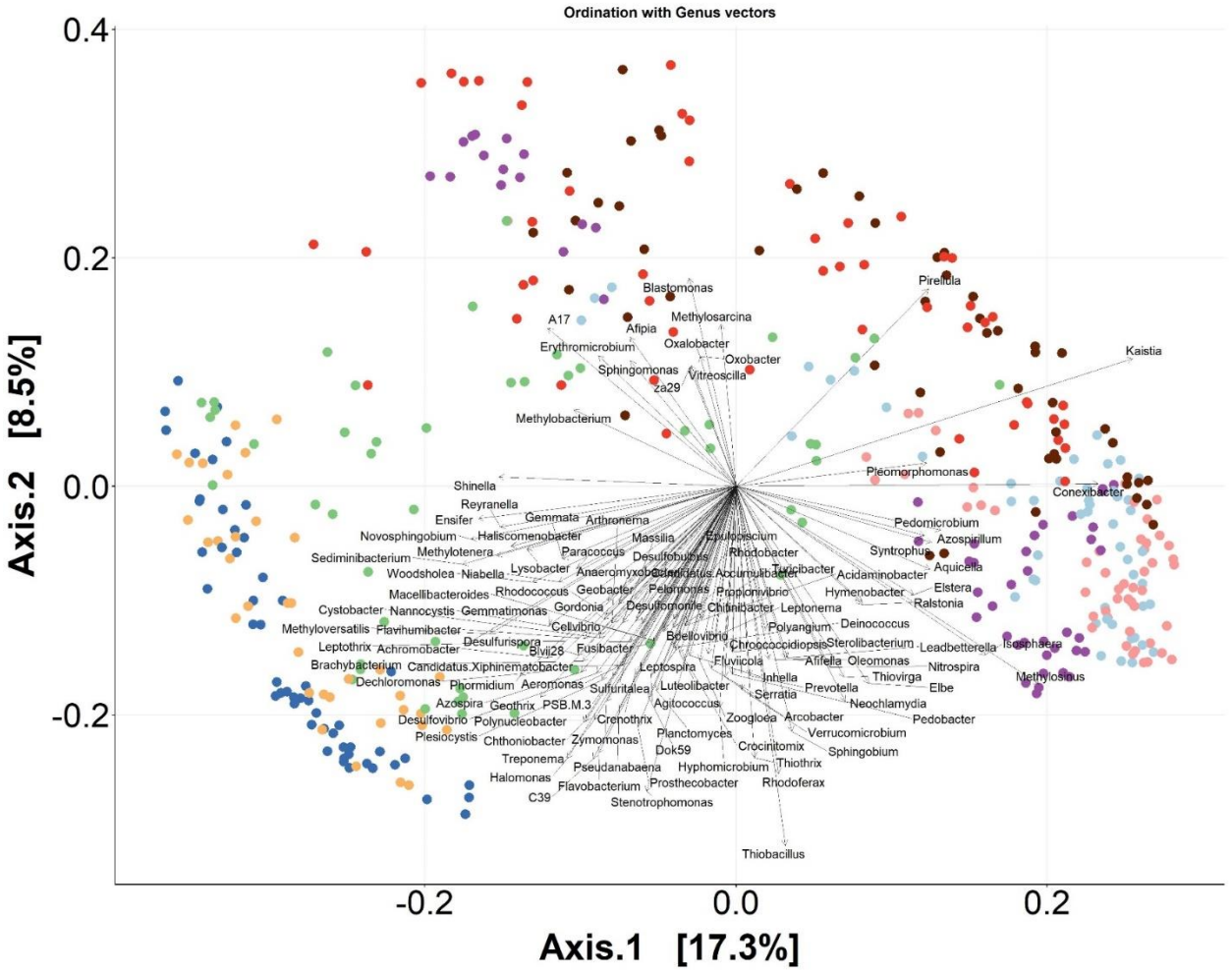
Anosim.UV.YR_SampleEvent	0.000999001	1000	0.9453
Anosim_Unweighted_3500	signif	permutations	statistic
Anosim.UV.YR_SampleEvent	0.000999001	1000	0.8324

SI Table F.18: UV Adonis Results

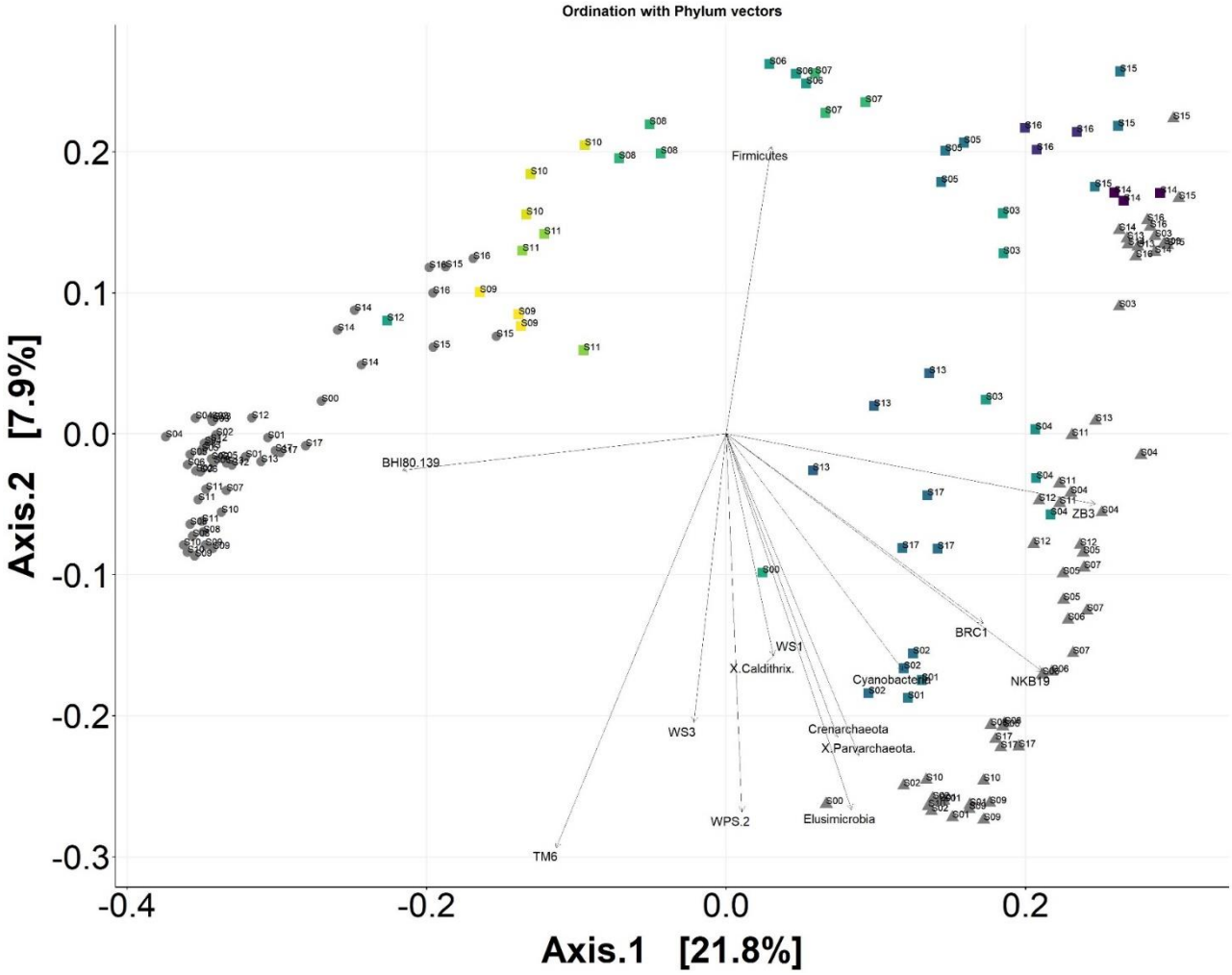
Adonis_Weighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.UV.YR_BackwashProximity_Skid	1	0.0995	0.0995	2.3696	0.0470	0.0228
Adonis.UV.YR_Temperature	1	0.1798	0.1798	4.4592	0.0850	0.0004
Adonis.UV.YR_Time_Days	1	0.2410	0.2410	6.1727	0.1139	0.0001
Adonis_Unweighted_3500	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Adonis.UV.YR_BackwashProximity_Skid	1	0.5360	0.5360	2.0723	0.0414	0.0228
Adonis.UV.YR_Temperature	1	0.8760	0.8760	3.4823	0.0676	0.0009
Adonis.UV.YR_Time_Days	1	1.5143	1.5143	6.3558	0.1169	0.0001



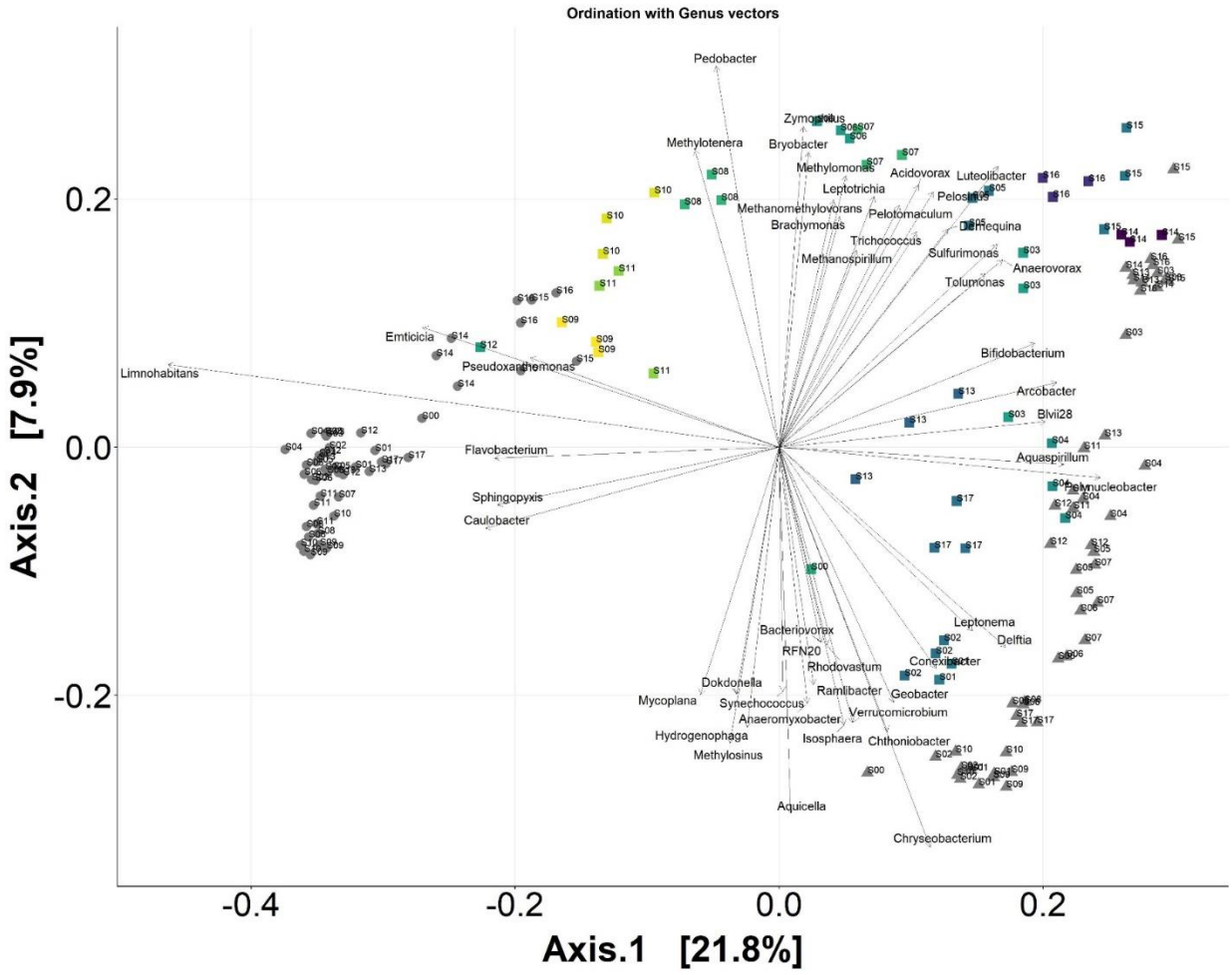
SI Figure F.1: UniFrac beta diversity plot comparing effluent samples from each indicated stage of treatment across the AWT pilot based on 16S rRNA gene amplicon sequencing with correspondence analysis at the Phylum classification.



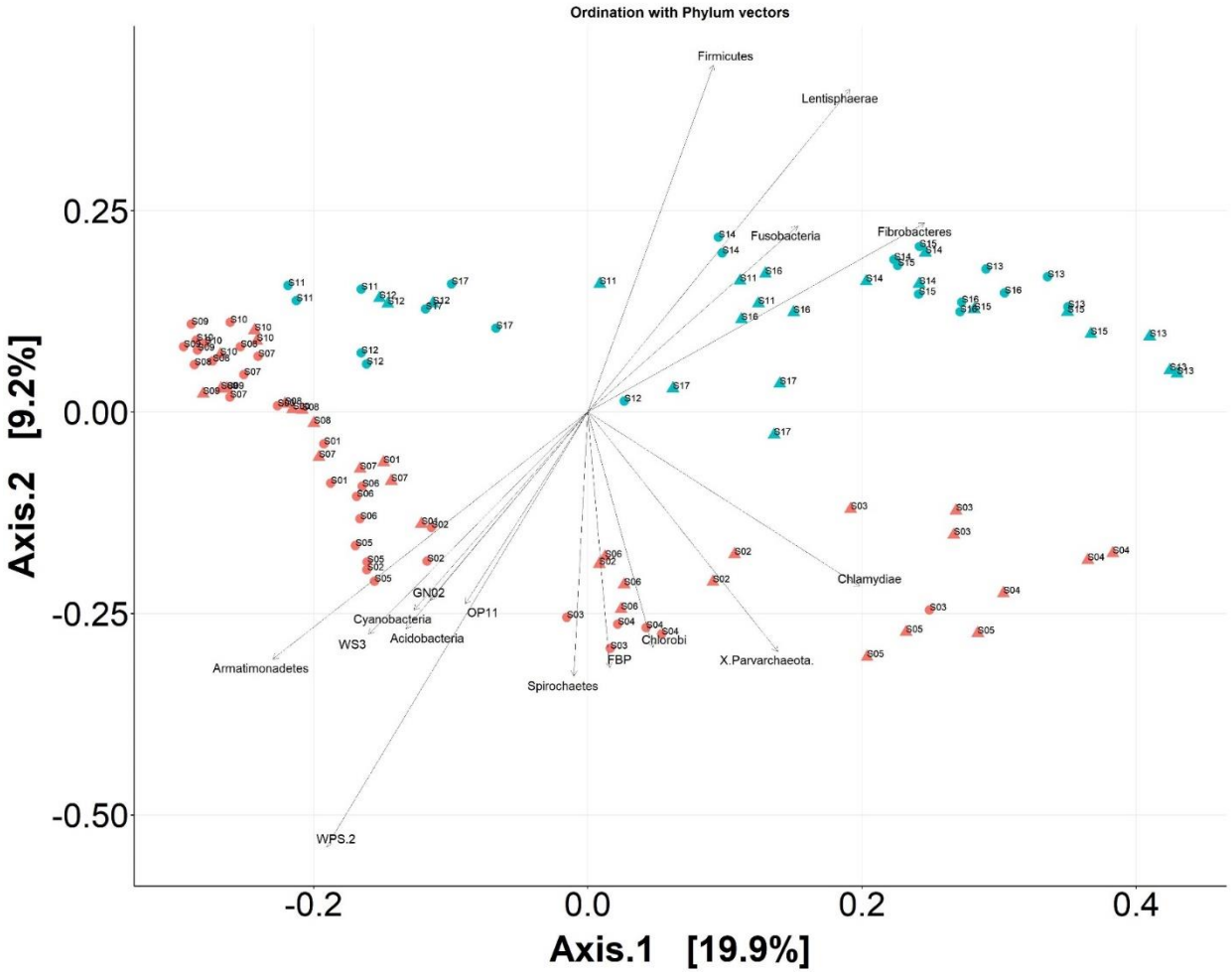
SI Figure F.2: UniFrac beta diversity plot comparing effluent samples from each indicated stage of treatment across the AWT pilot based on 16S rRNA gene amplicon sequencing with correspondence analysis at the Genus classification.



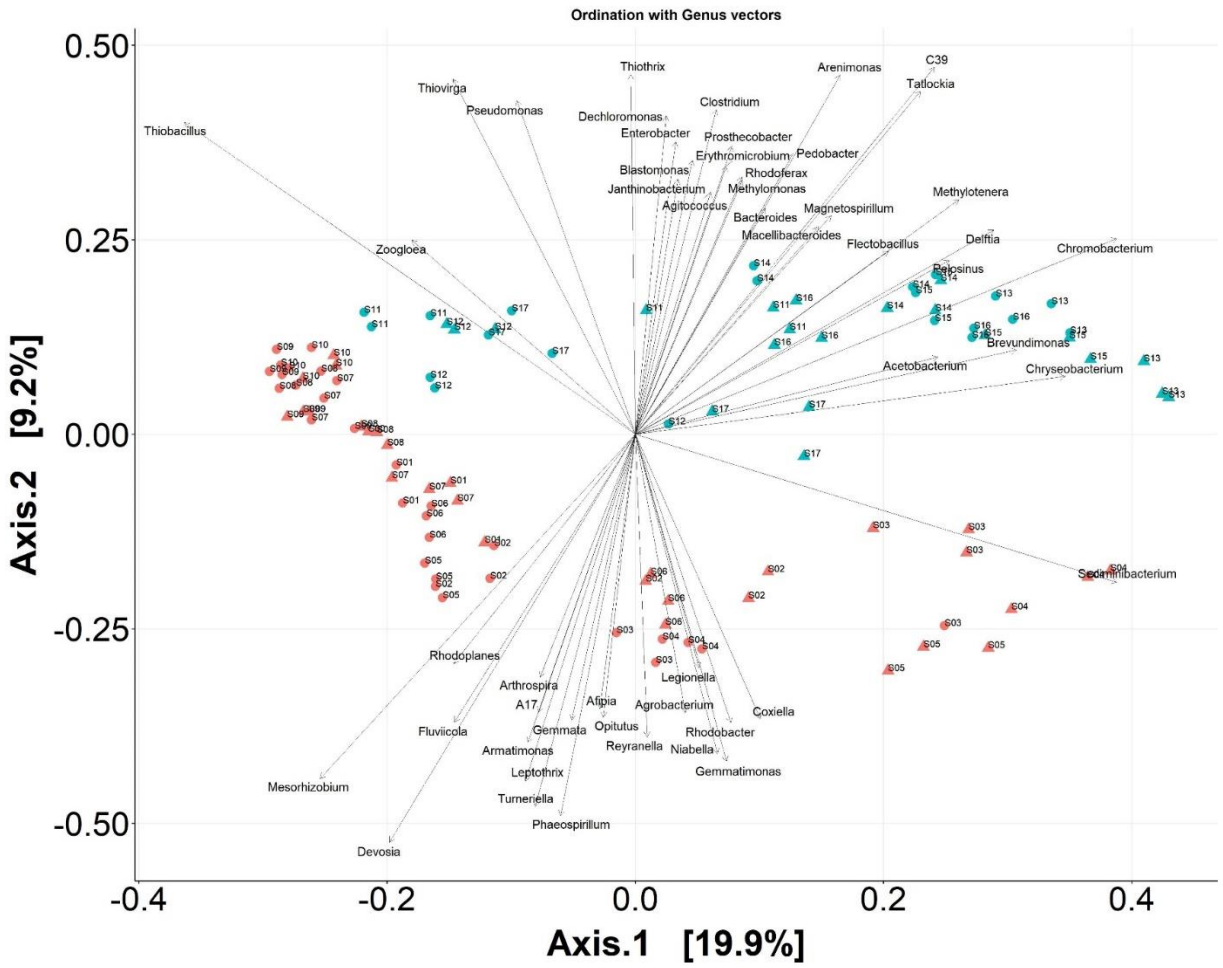
SI Figure F.3: Unweighted UniFrac beta diversity plot of ozone effluent microbial communities based on 16S rRNA gene amplicon sequencing. Correspondence analysis conducted at the Phylum classification level.



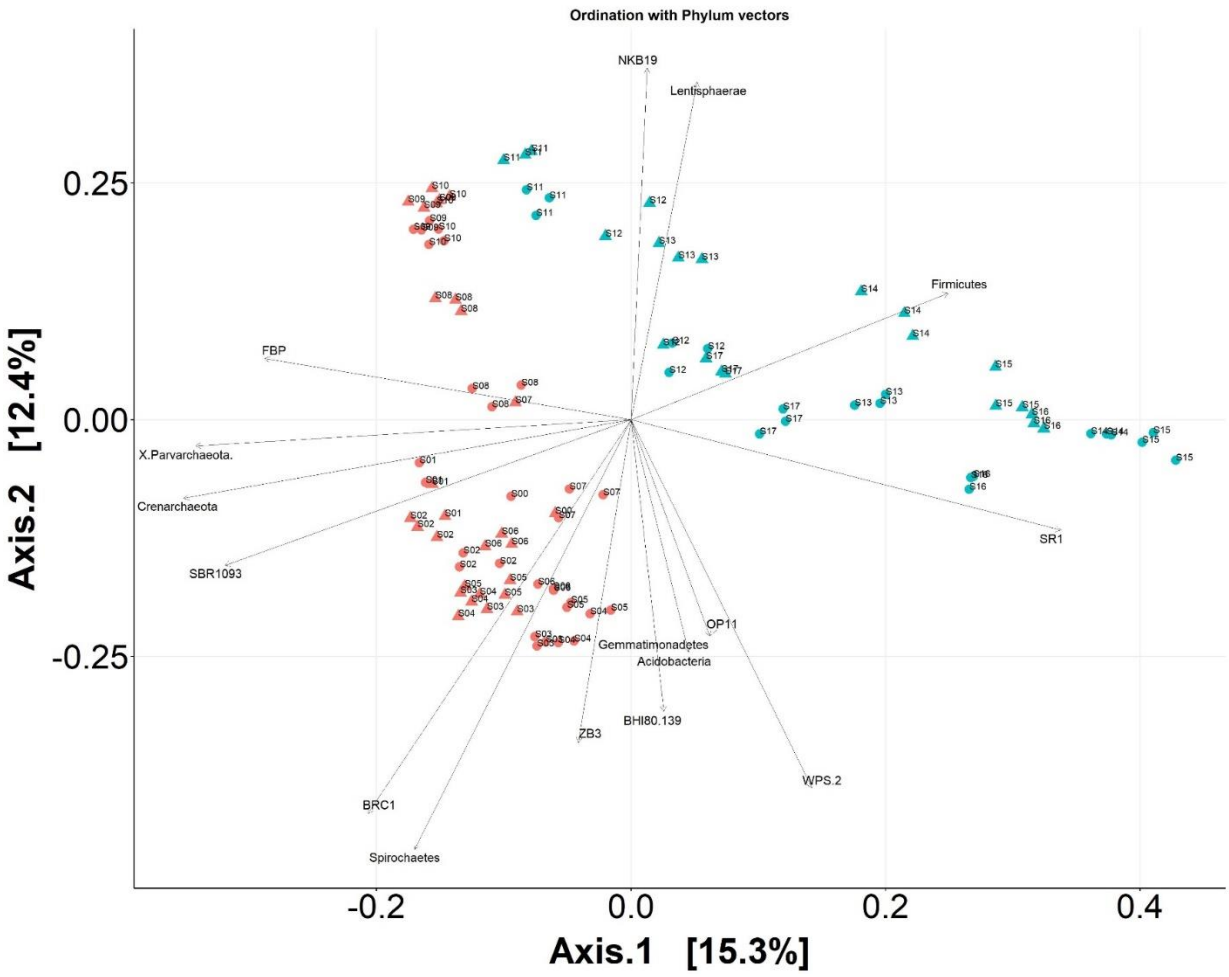
SI Figure F.4: Unweighted UniFrac beta diversity plot of ozone effluent microbial communities based on 16S rRNA gene amplicon sequencing. Correspondence analysis conducted at the Genus classification level.



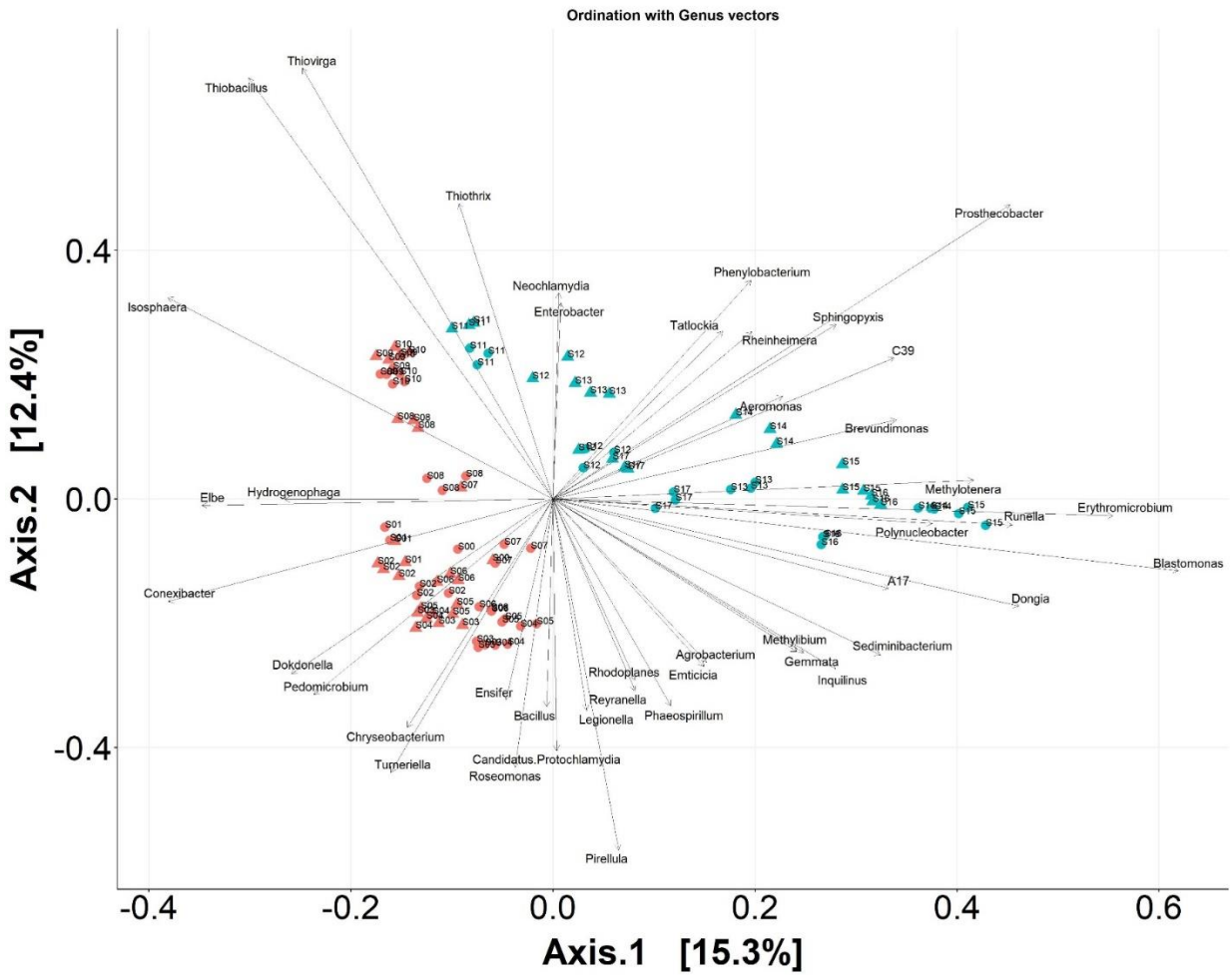
SI Figure F.5: Unweighted UniFrac beta diversity plot for both BAC5 and BAC10 effluent samples with sample event identified. Correspondence analysis conducted at the Phylum classification level.



SI Figure F.6: Unweighted UniFrac beta diversity plot for both BAC5 and BAC10 effluent samples with sample event identified. Correspondence analysis conducted at the Genus classification level.

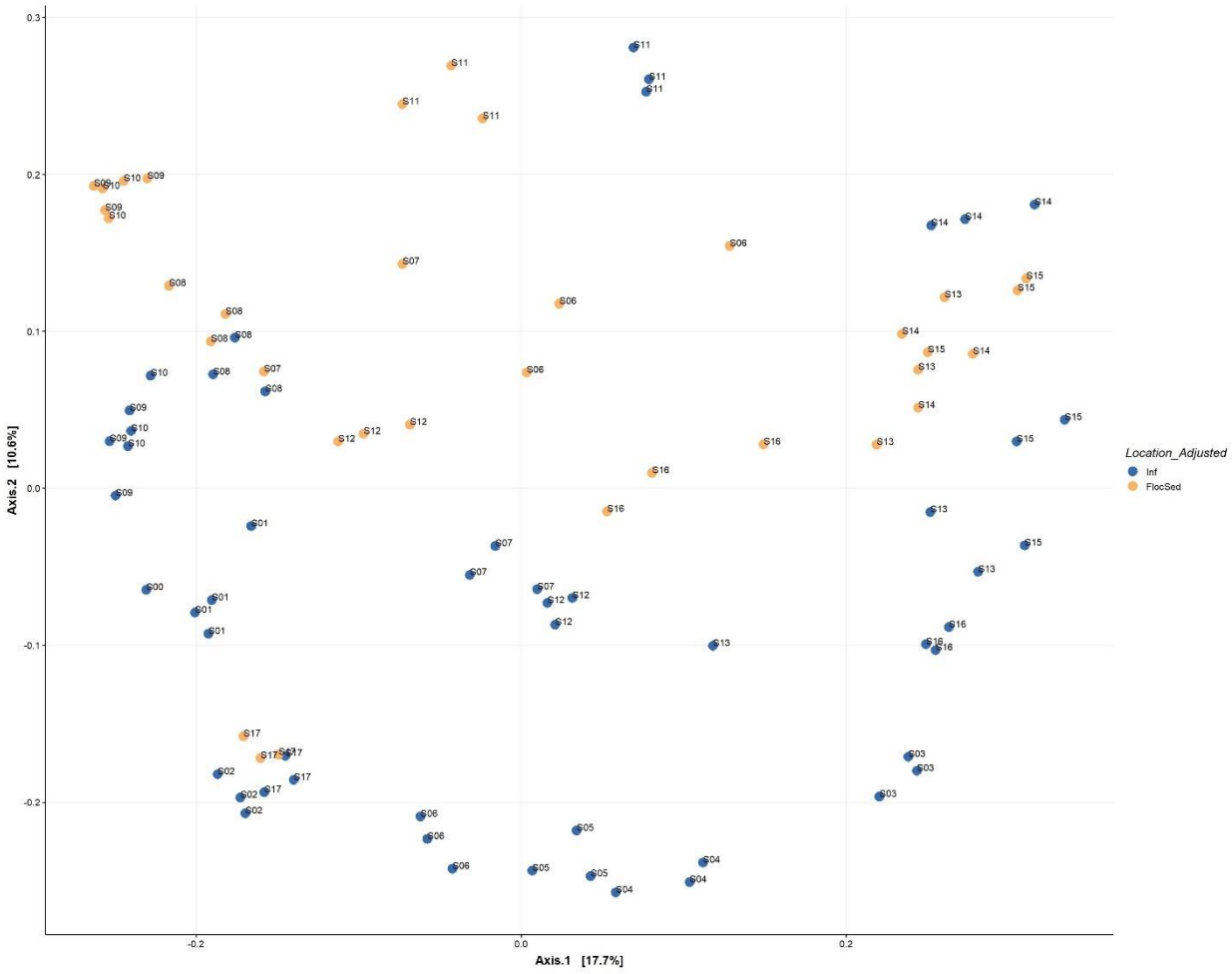


SI Figure F.7: Unweighted UniFrac beta diversity plot for both GAC10 and GAC20 effluent samples with sample event identified. Correspondence analysis conducted at the Phylum classification level.

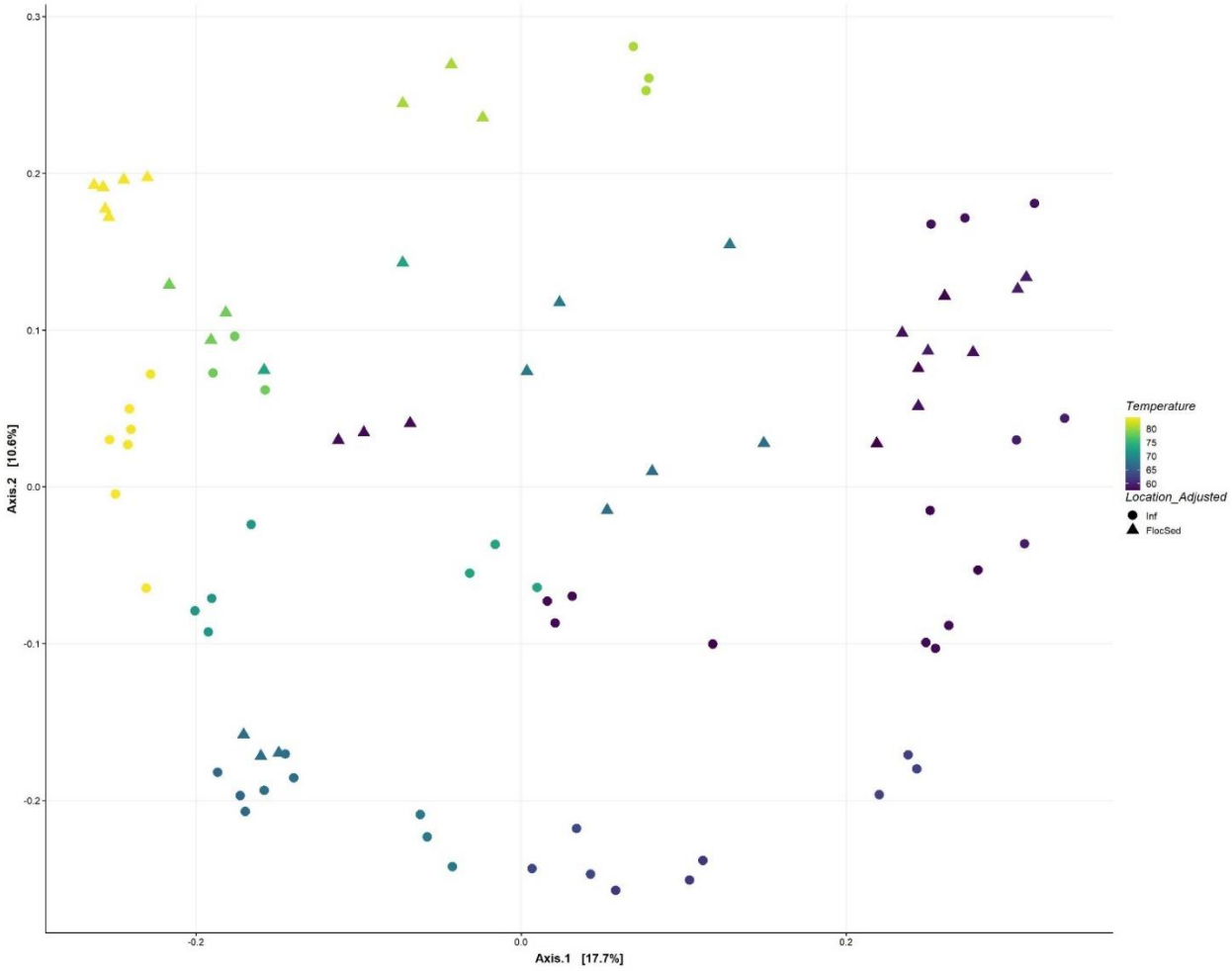


SI Figure F.8: Unweighted UniFrac beta diversity plot for both GAC10 and GAC20 effluent samples with sample event identified. Correspondence analysis conducted at the Genus classification level.

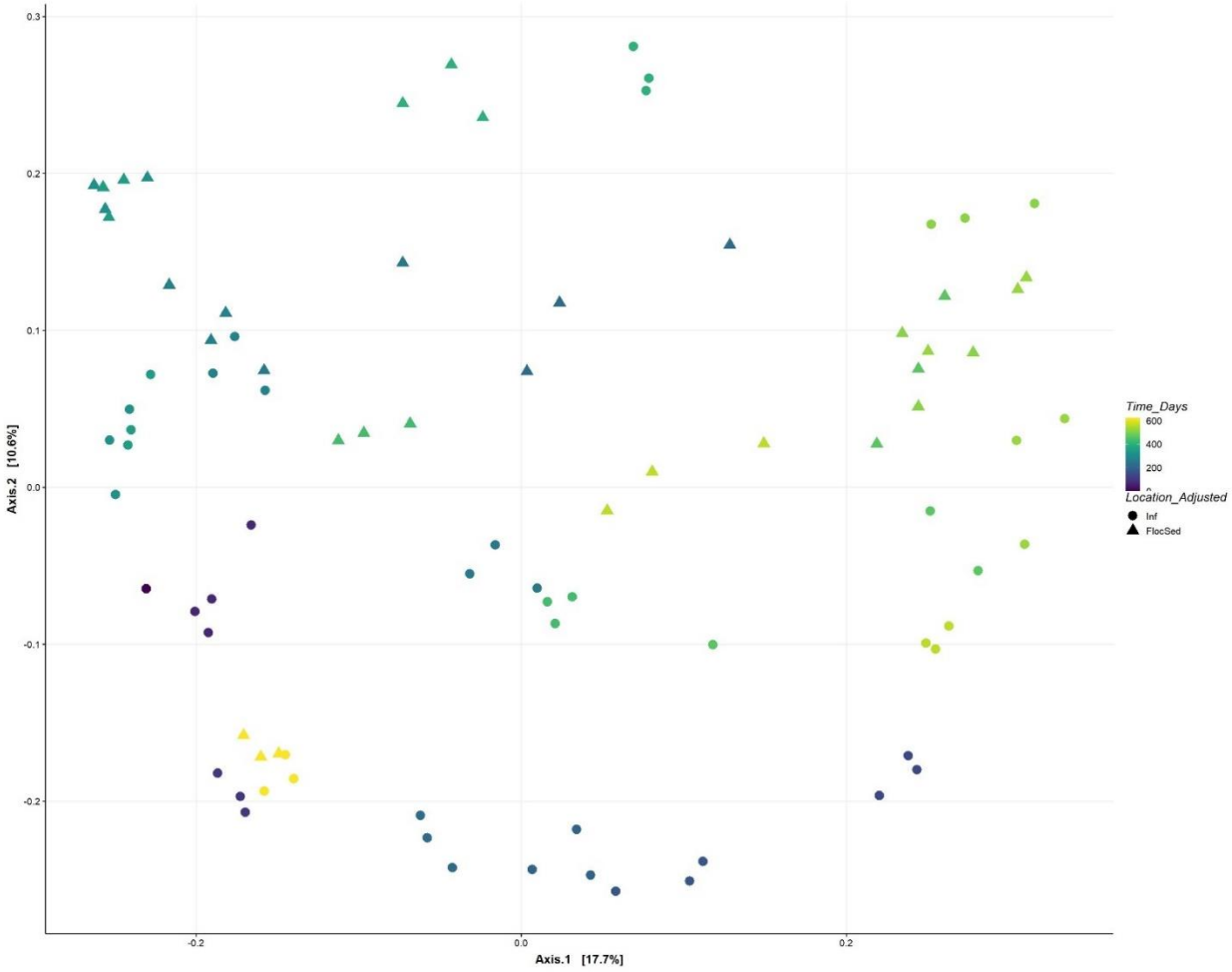
Subsection 1: Supplemental Ordination Plots: Influent and FlocSed



SI Figure F.9: Unweighted UniFrac beta diversity plot with sample event identified.

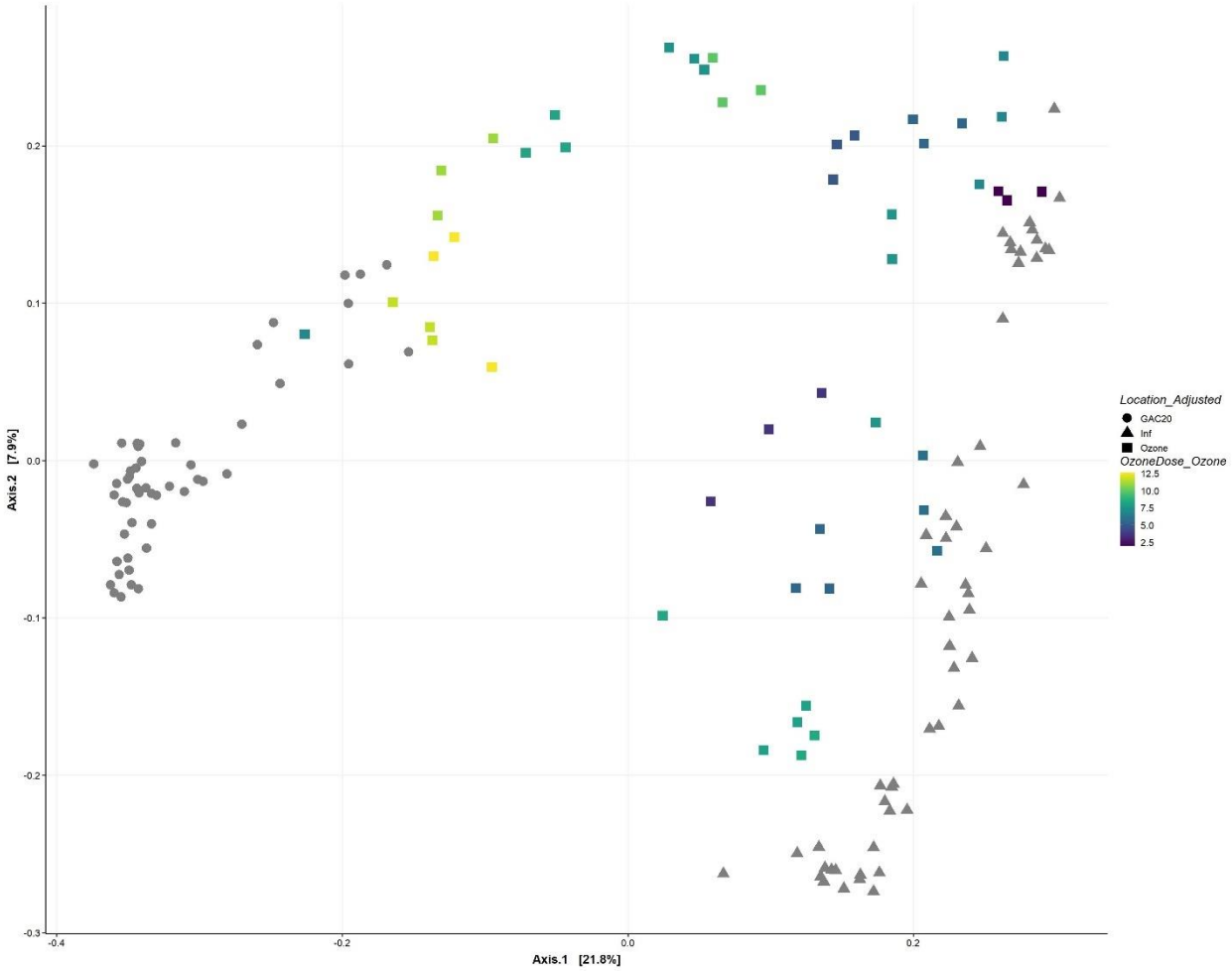


SI Figure F.10 Unweighted UniFrac beta diversity plot with temperature identified.

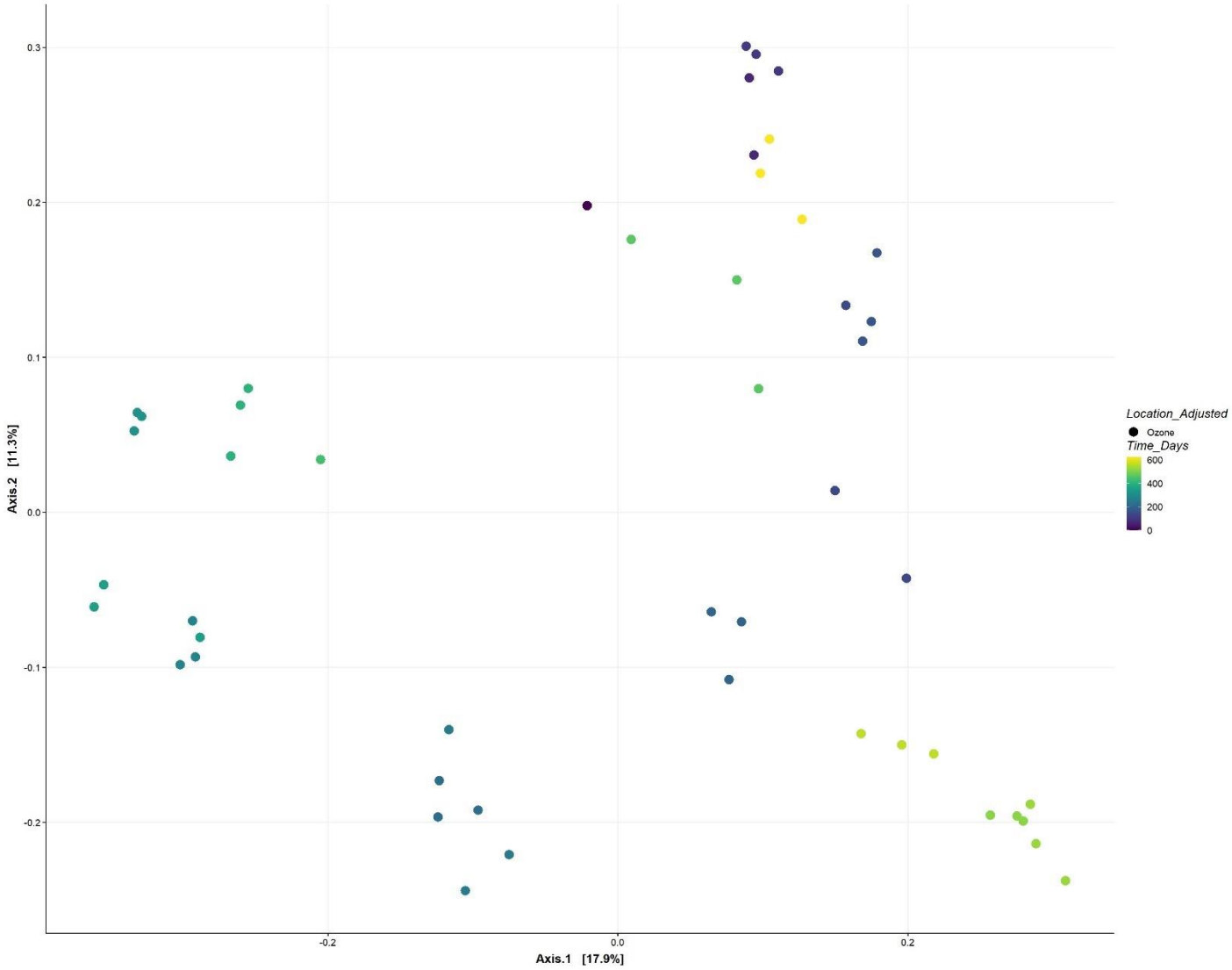


SI Figure F.11: Unweighted UniFrac beta diversity plot with pilot runtime identified.

Subsection 2: Supplemental Ordination Plots: Ozonation

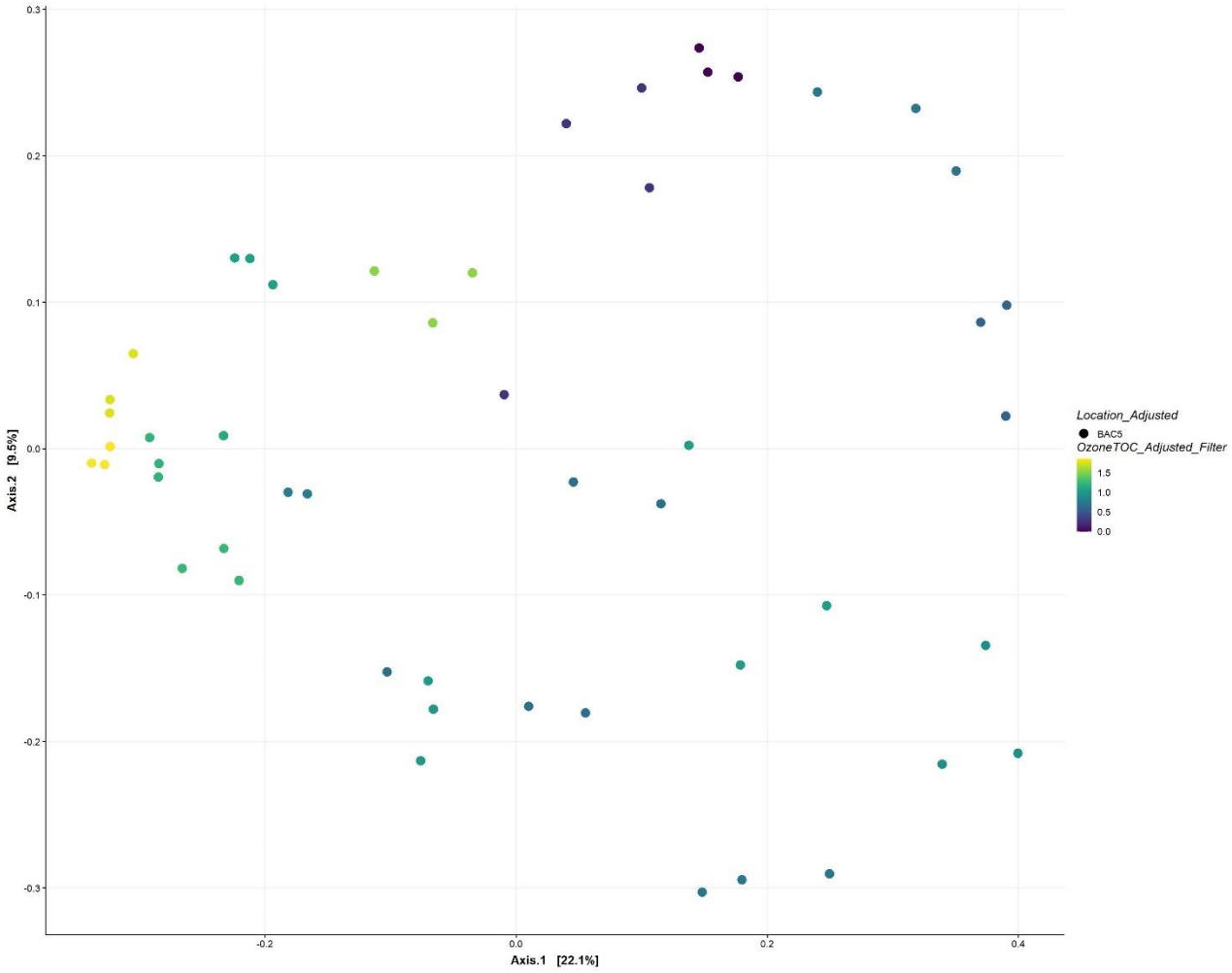


SI Figure F.12: Unweighted UniFrac beta diversity plot with ozone dose identified.

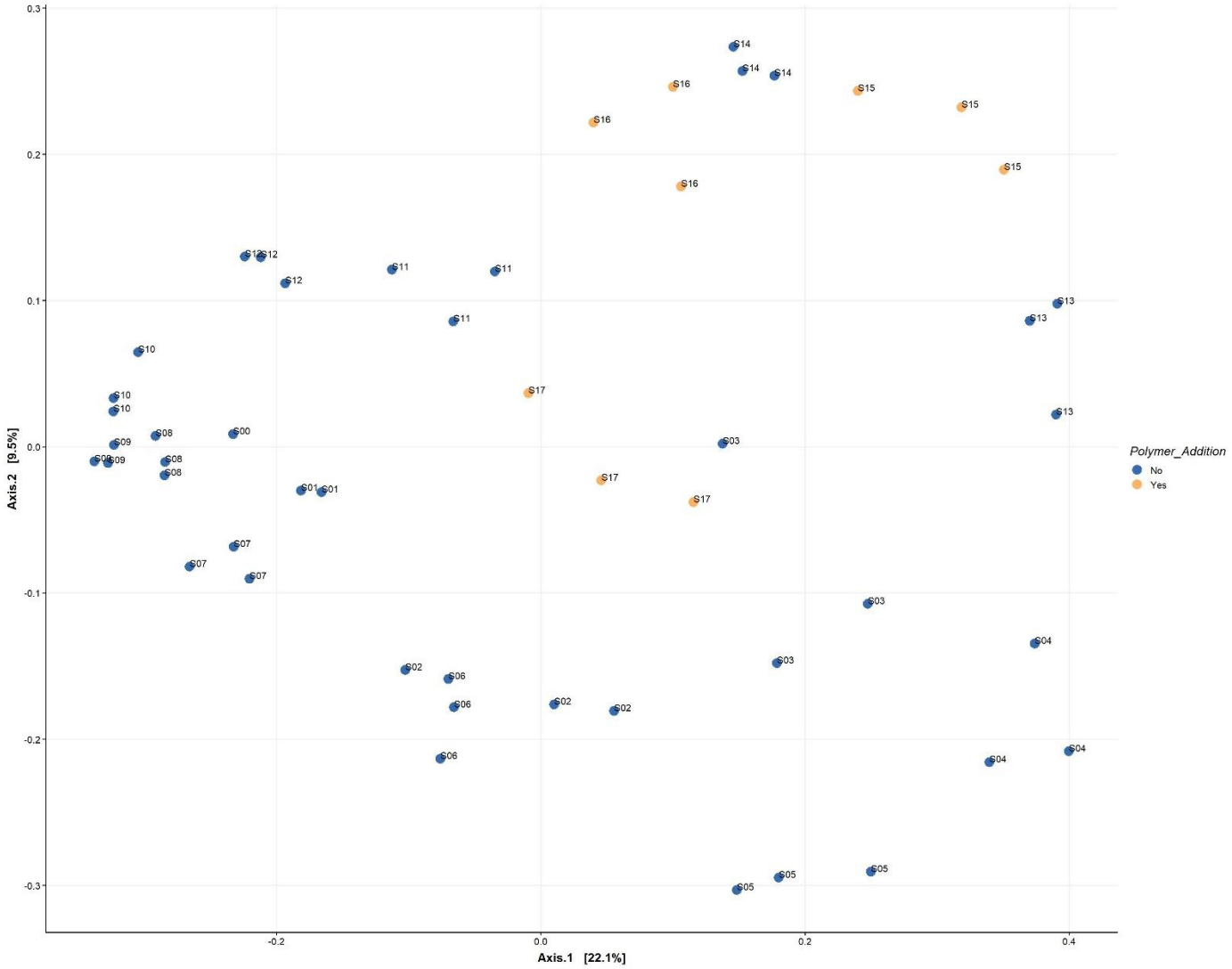


SI Figure F.13: Unweighted UniFrac beta diversity plot with pilot runtime identified.

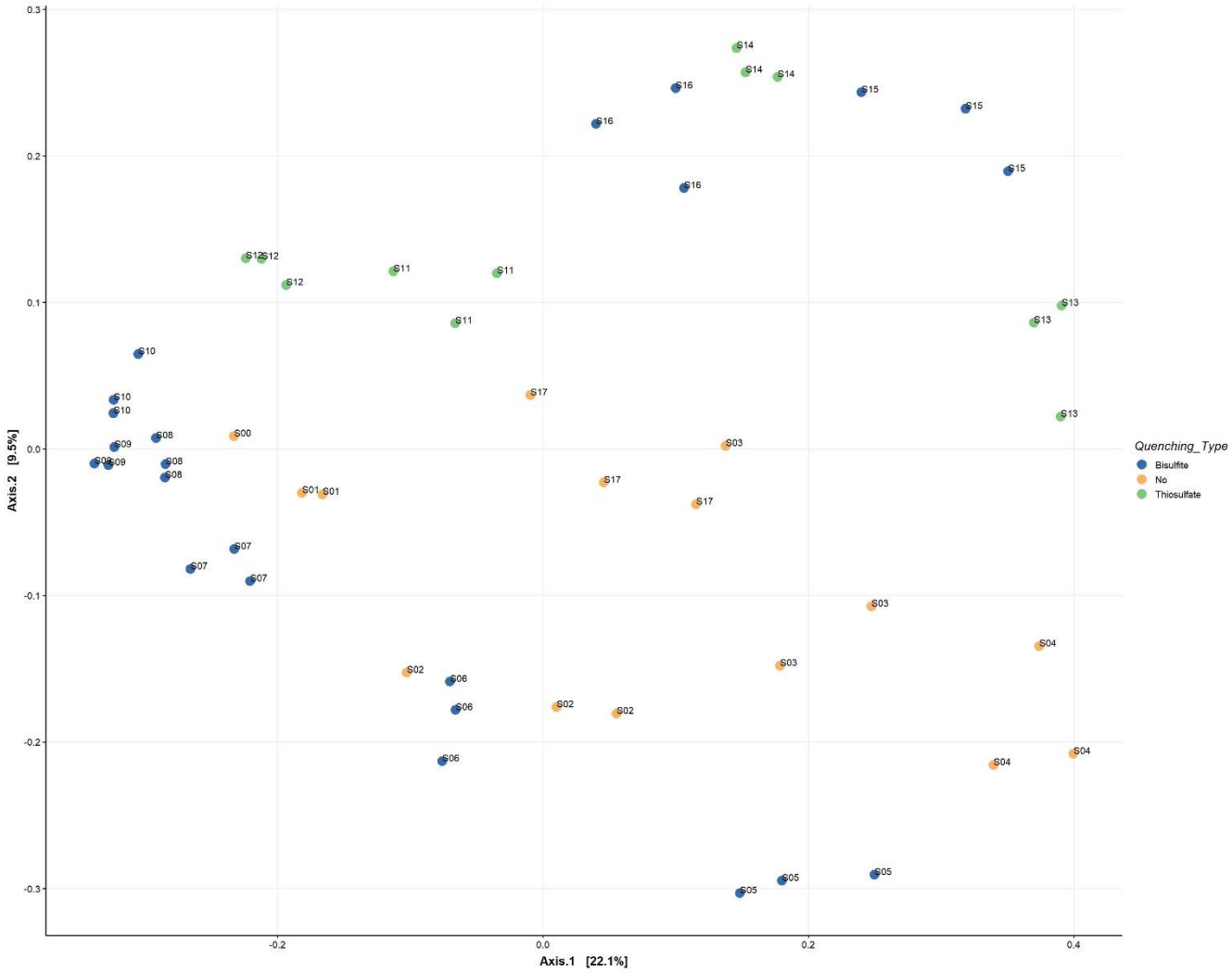
Subsection 3: Supplemental Ordination Plots: BAC Filtration



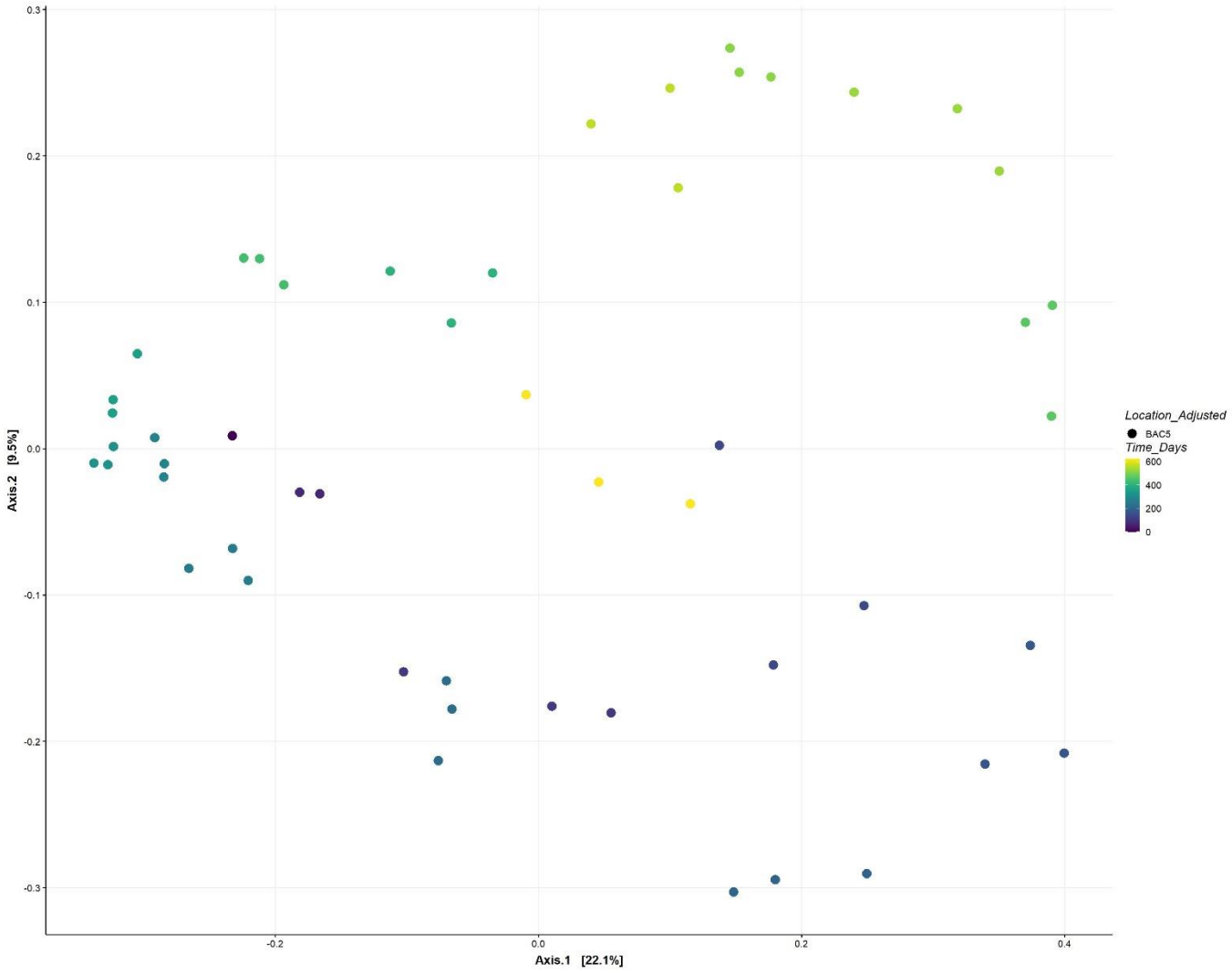
SI Figure F.14: Unweighted UniFrac beta diversity plot for BAC5 with adjusted O_3:TOC ratio identified.



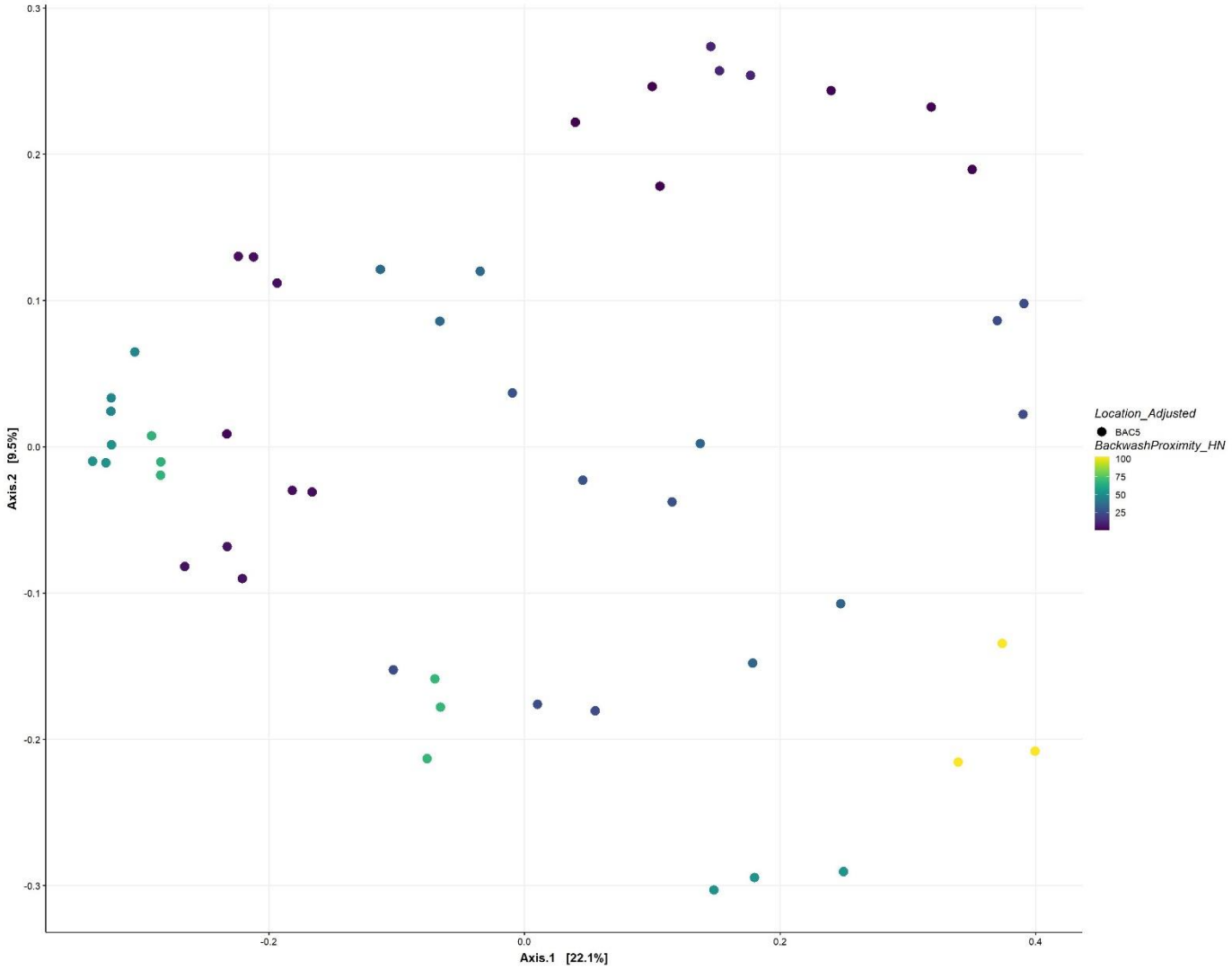
SI Figure F.15: Unweighted UniFrac beta diversity plot for BAC5 with polymer addition identified.



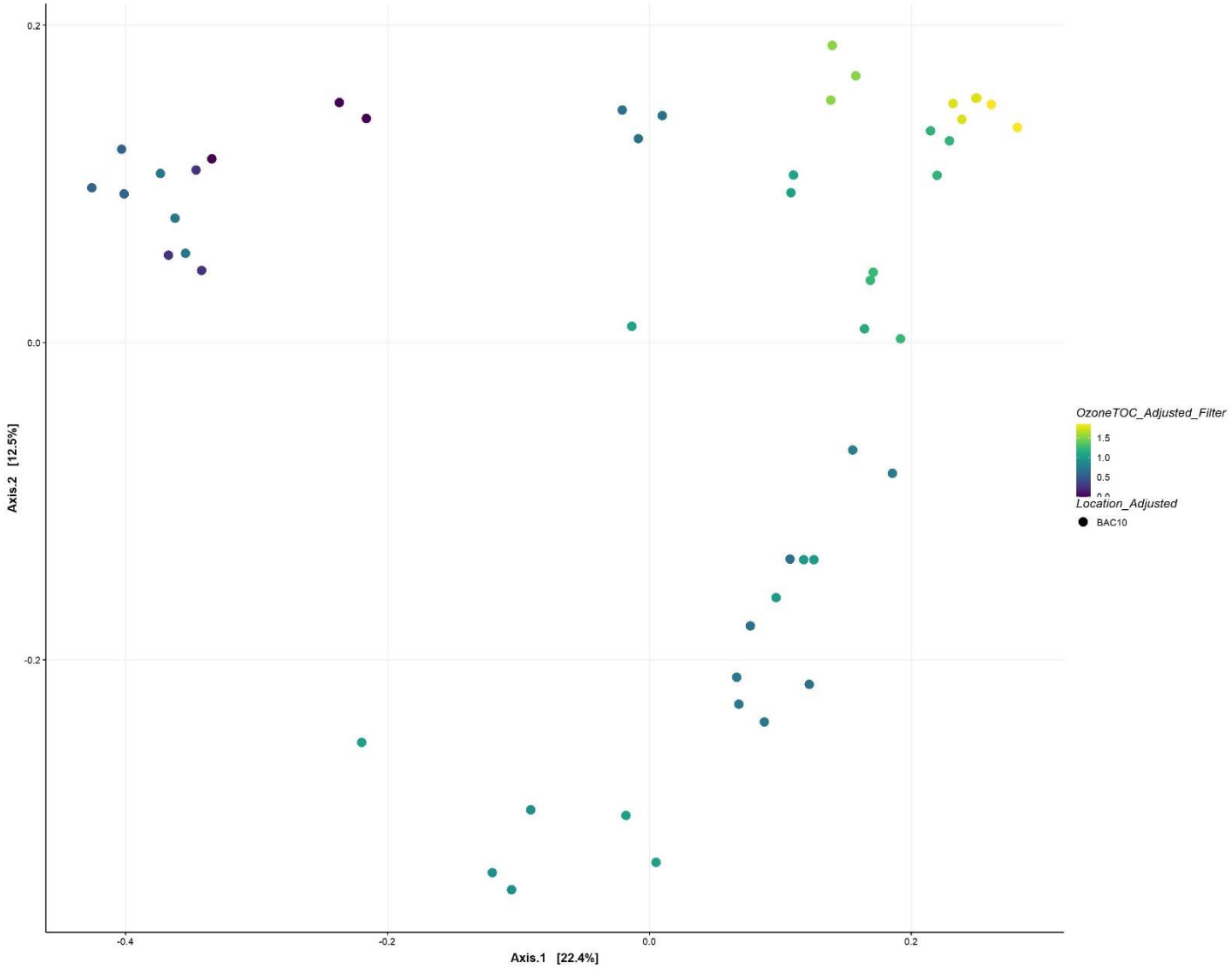
SI Figure F.16: Unweighted UniFrac beta diversity plot for BAC5 with quenching type identified.



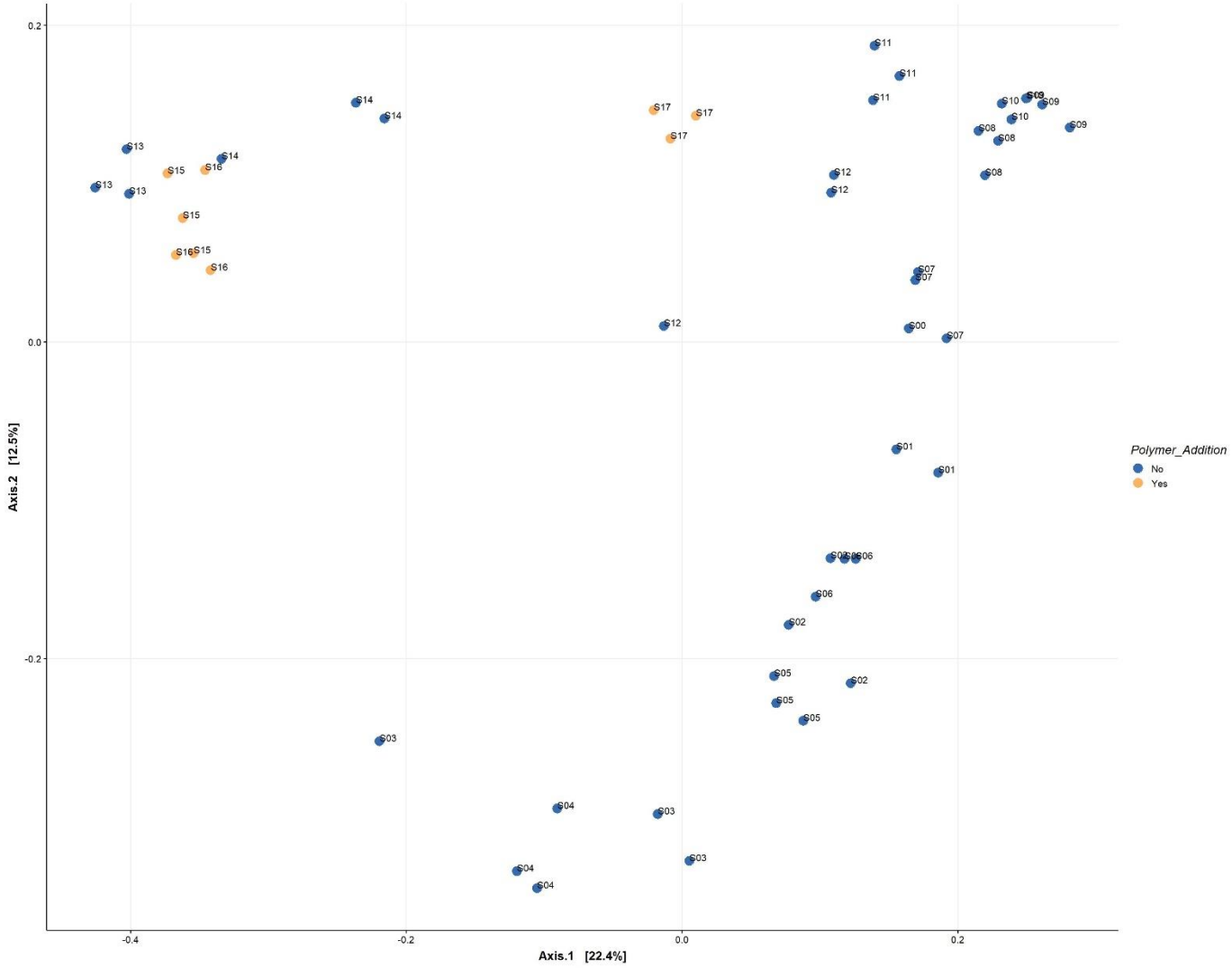
SI Figure F.17: Unweighted UniFrac beta diversity plot for BAC5 with pilot runtime identified.



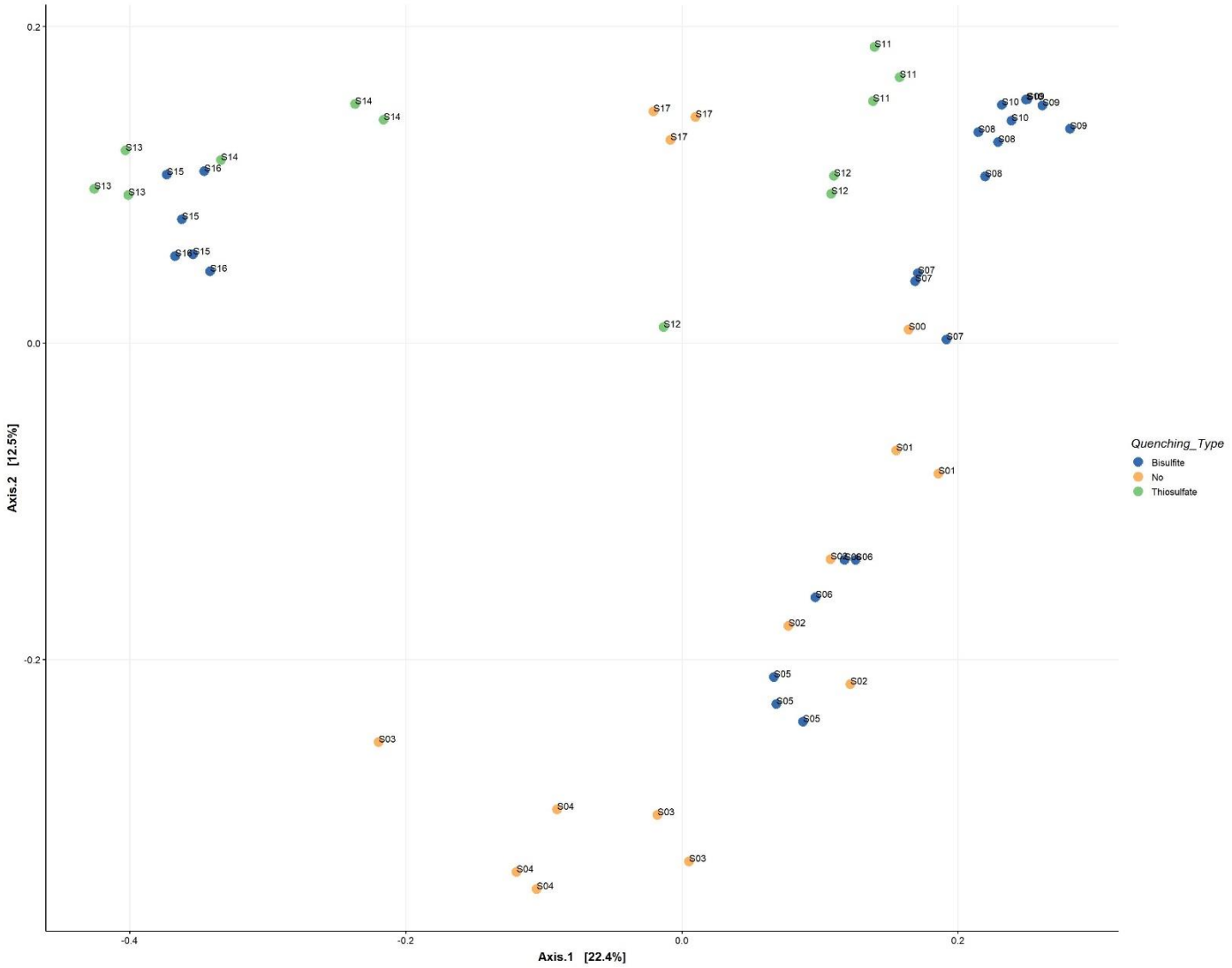
SI Figure F.18: Unweighted UniFrac beta diversity plot for BAC5 with temporal proximity to backwash identified.



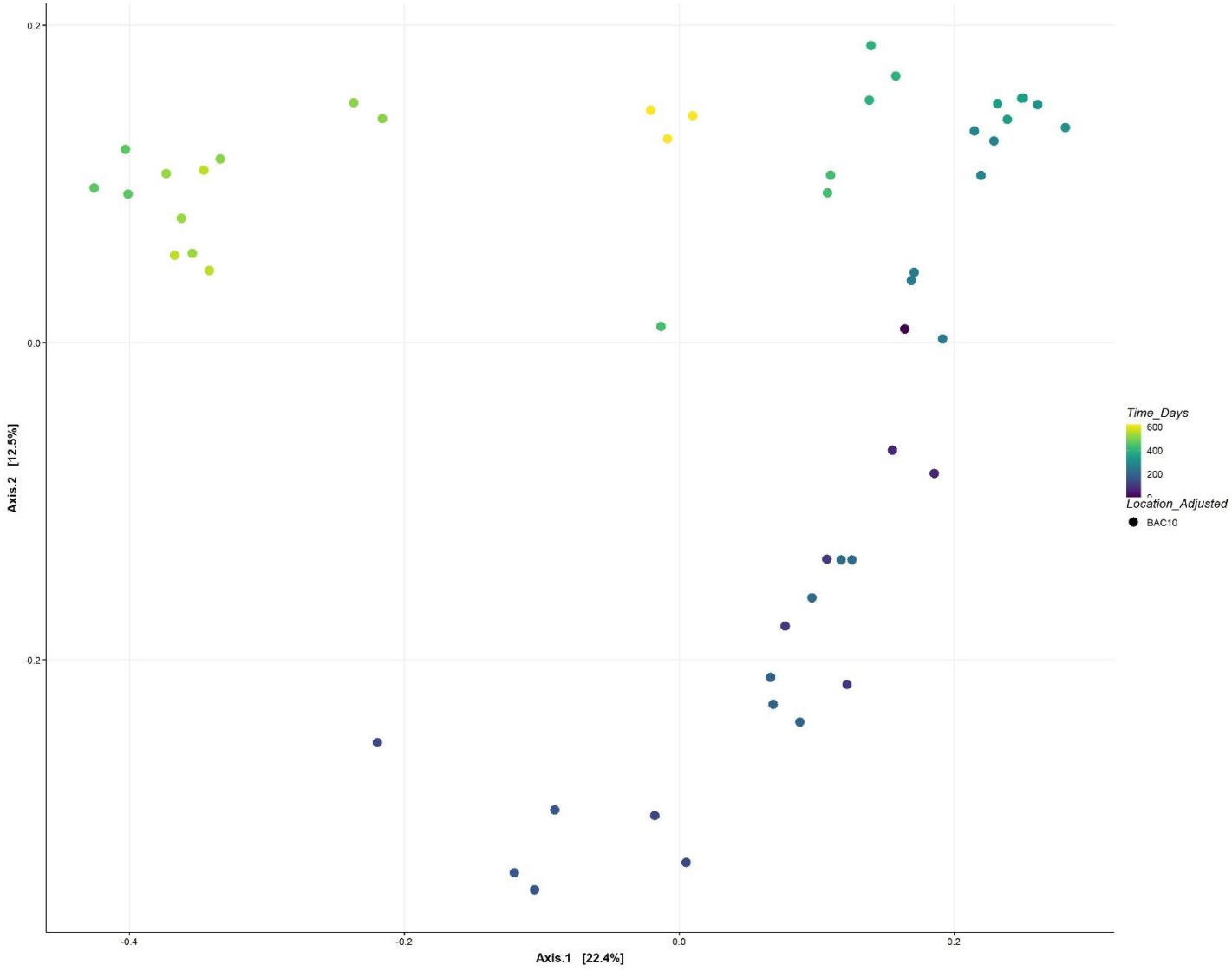
SI Figure F.19: Unweighted UniFrac beta diversity plot for BAC5 with adjusted O_3:TOC ratio identified.



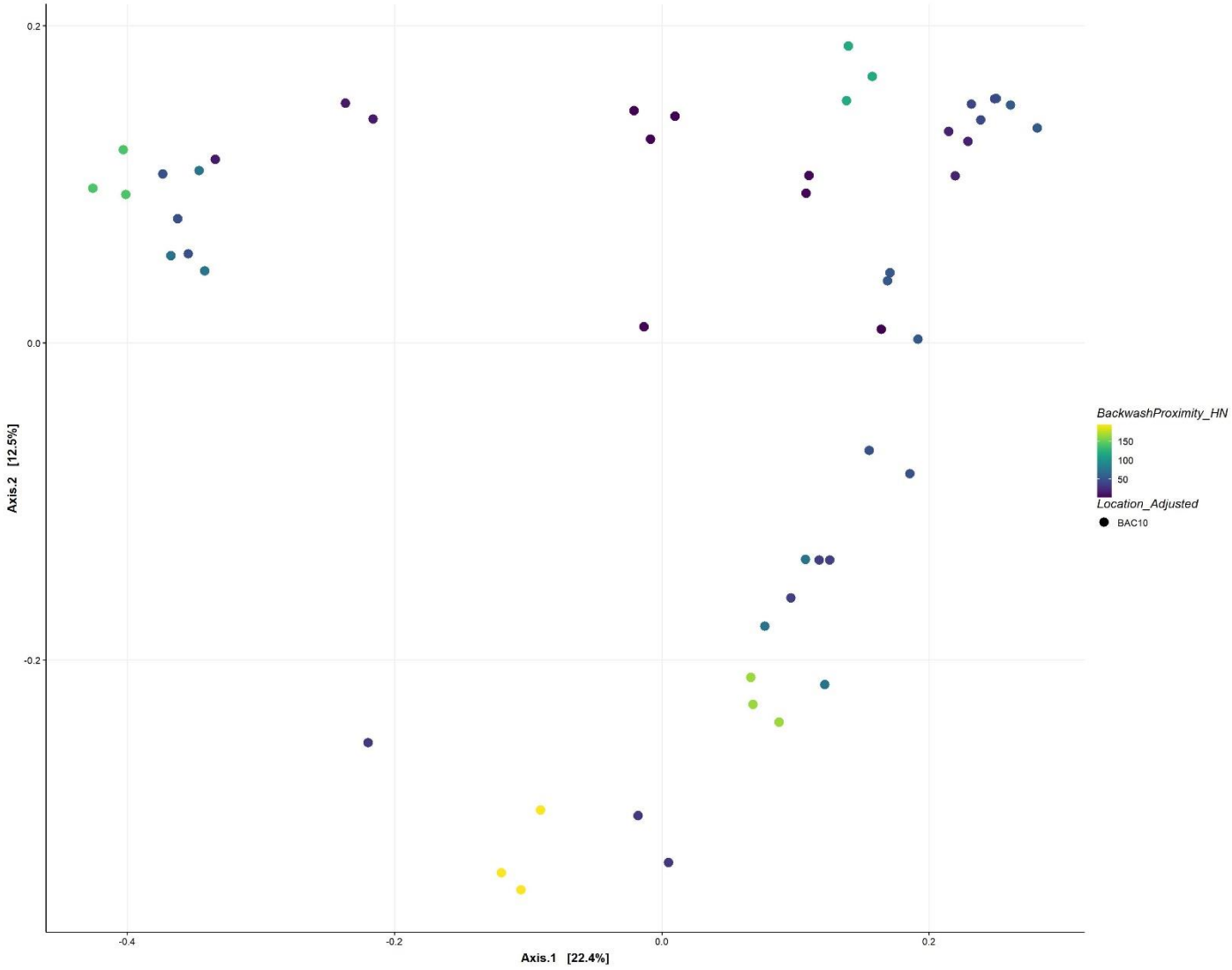
SI Figure F.20: Unweighted UniFrac beta diversity plot for BAC10 with polymer addition identified.



SI Figure F.21: Unweighted UniFrac beta diversity plot for BAC10 with quenching type identified.

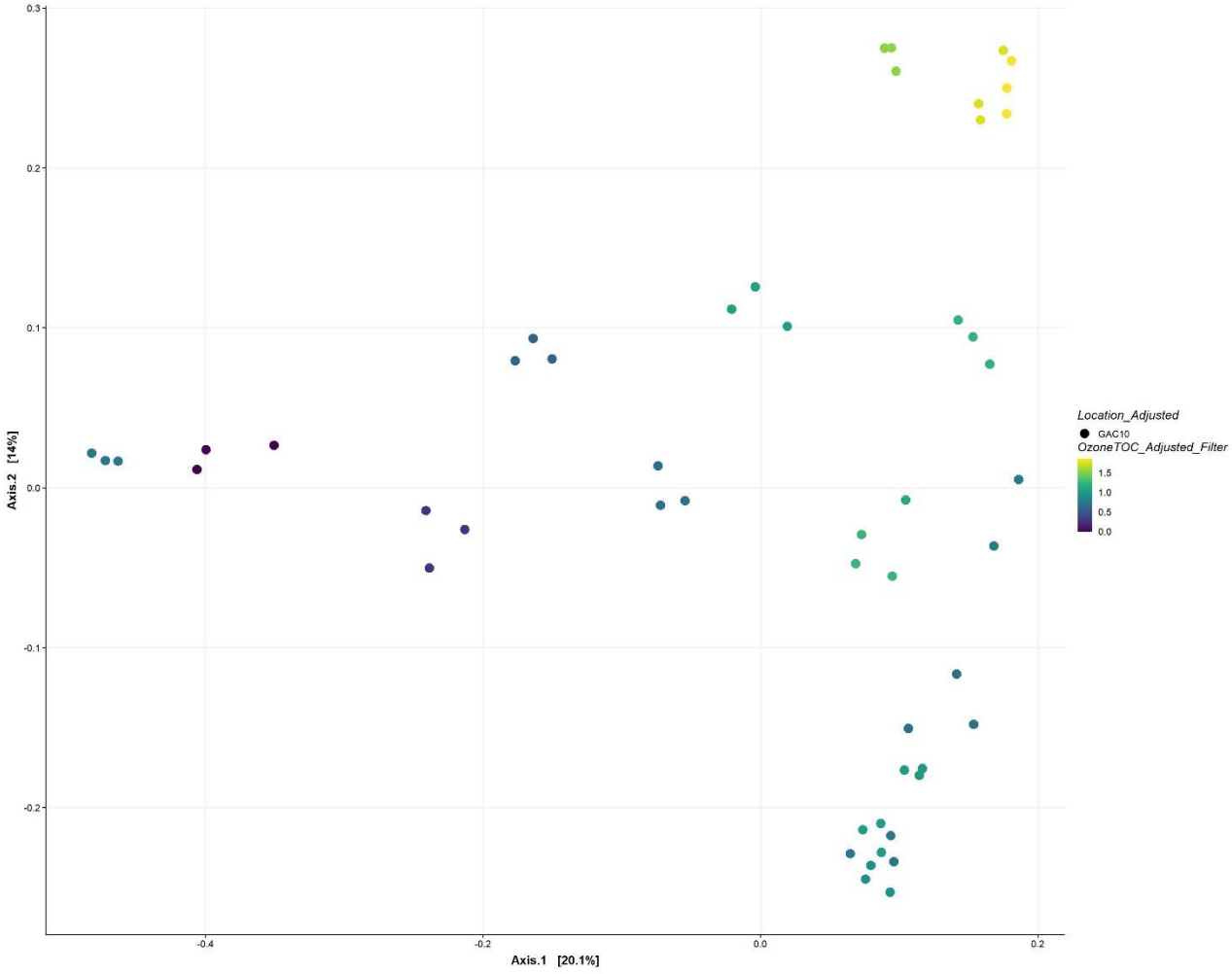


SI Figure F.22: Unweighted UniFrac beta diversity plot for BAC10 with pilot runtime identified.

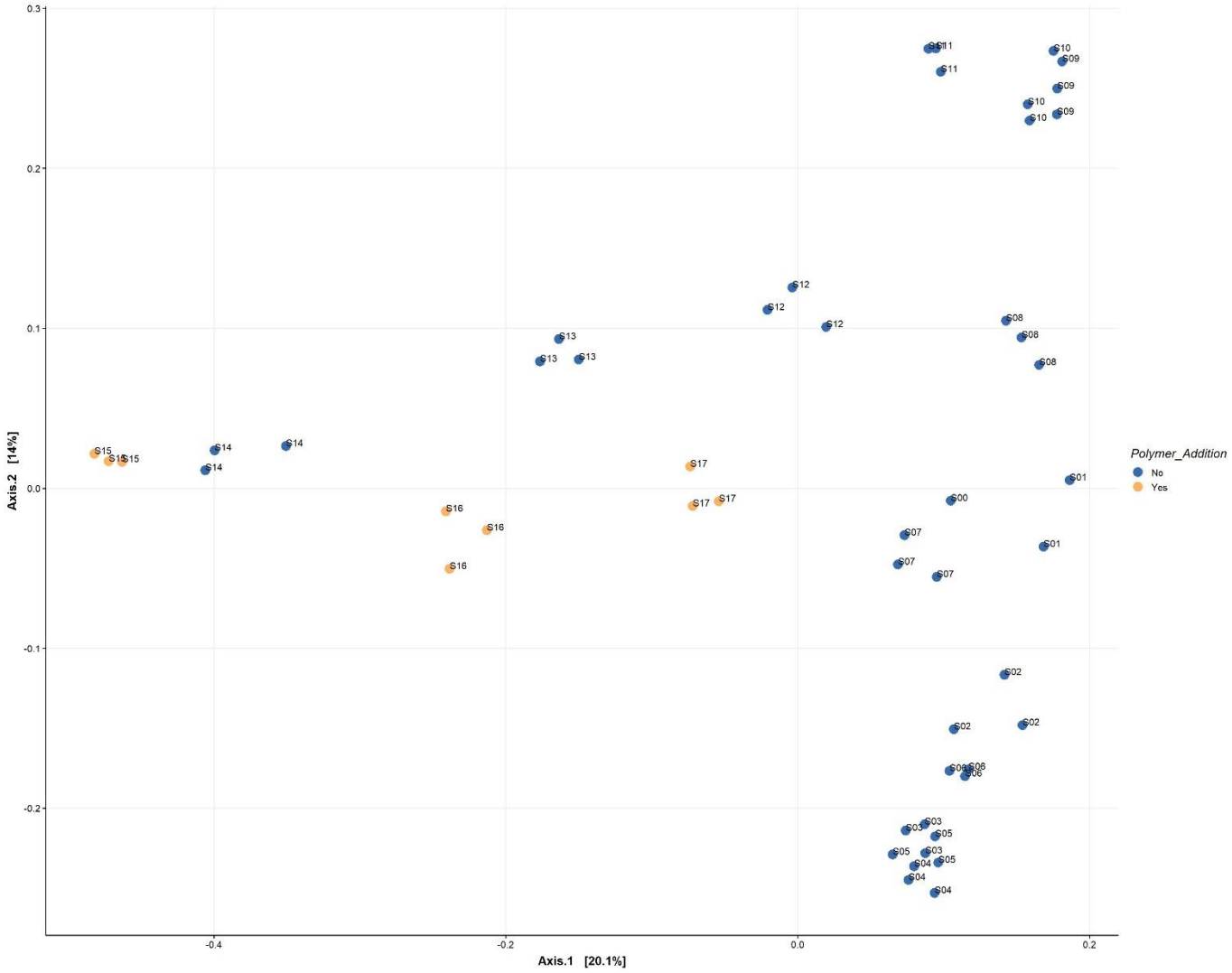


SI Figure F.23: Unweighted UniFrac beta diversity plot for BAC10 with temporal proximity to backwash identified.

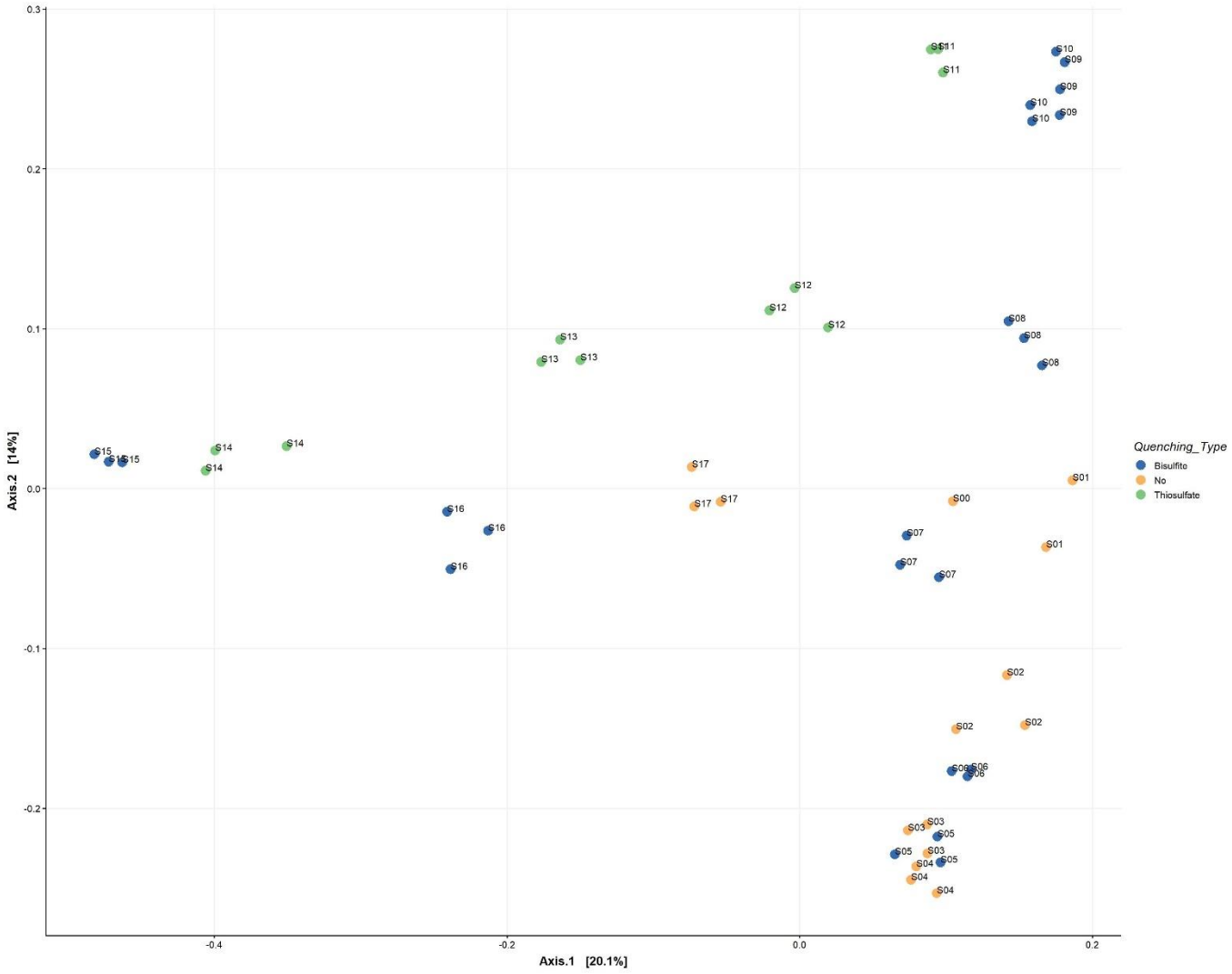
Subsection 4: Supplemental Ordination Plots: GAC Contacting



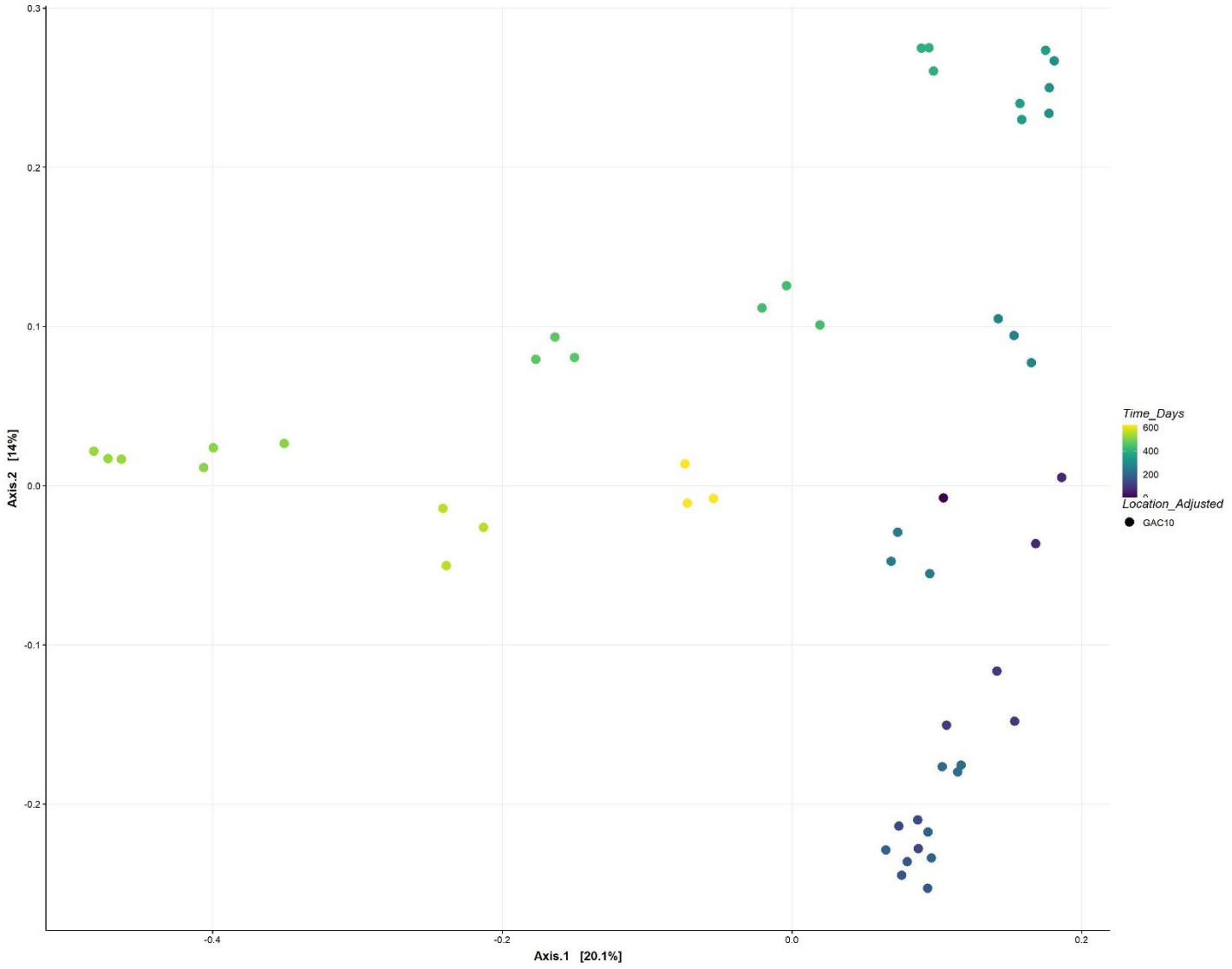
SI Figure F.24: Unweighted UniFrac beta diversity plot for GAC10 with adjusted O_3:TOC ratio identified.



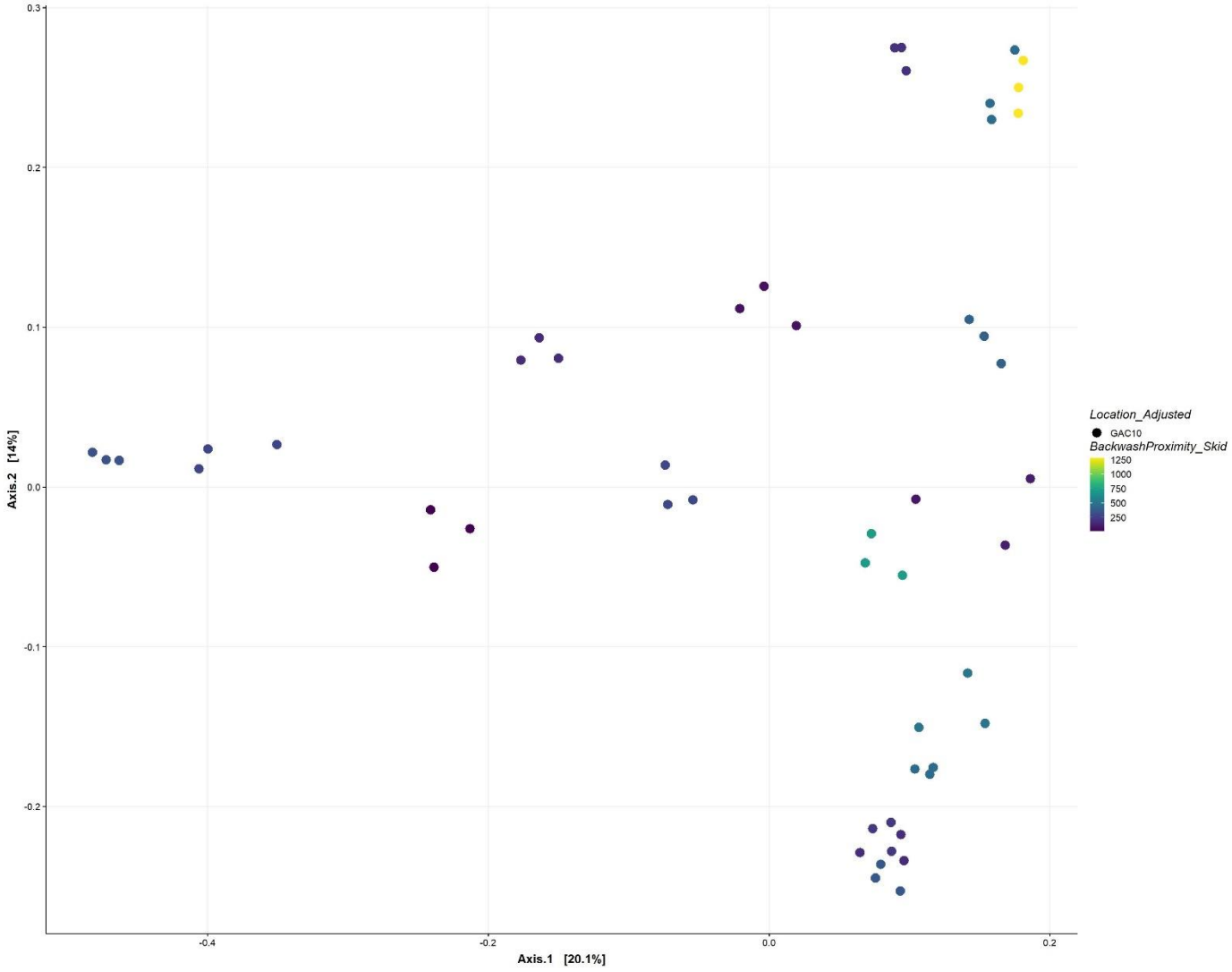
SI Figure F.25: Unweighted UniFrac beta diversity plot for GAC10 with polymer addition identified.



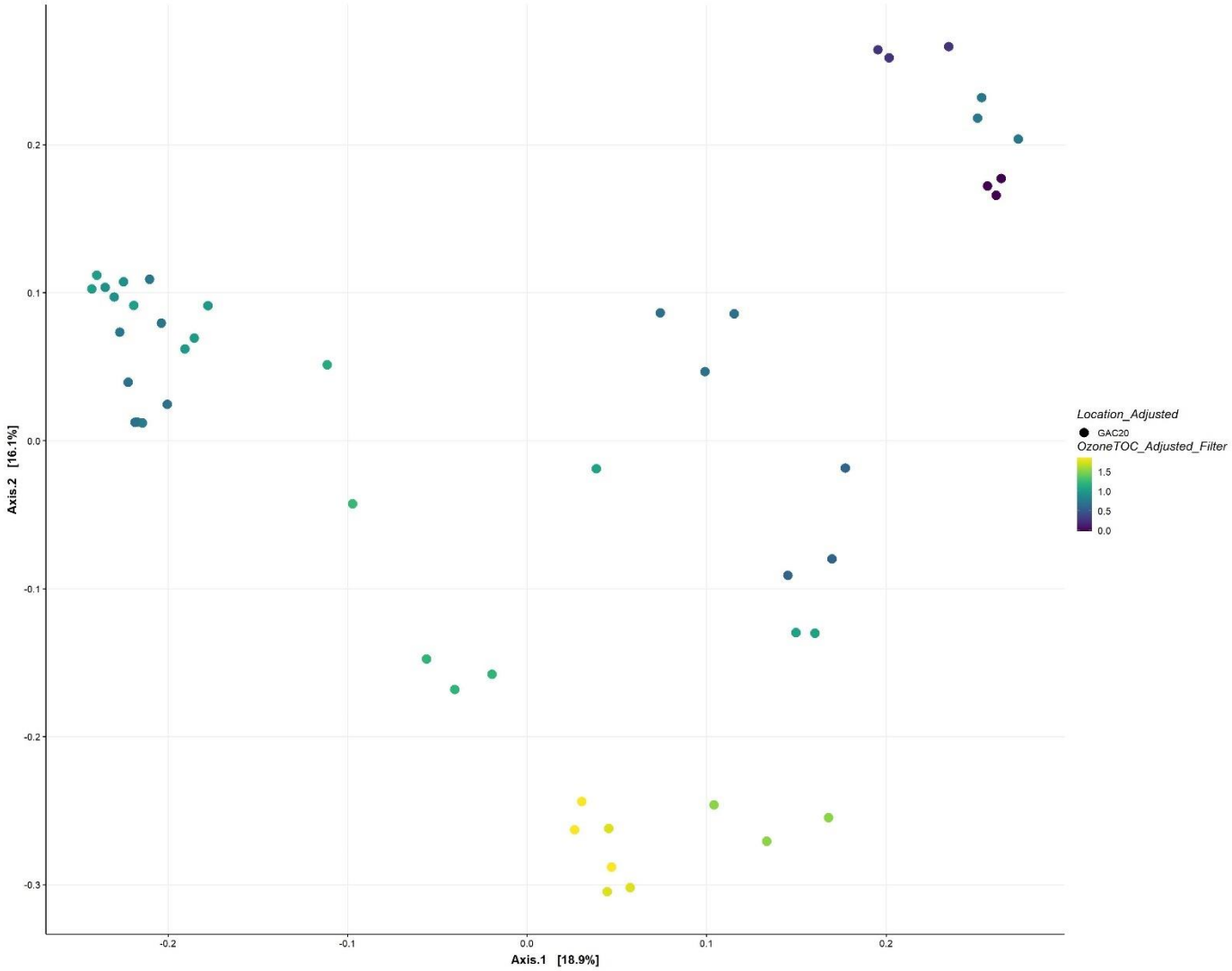
SI Figure F.26: Unweighted UniFrac beta diversity plot for GAC10 with quenching type identified.



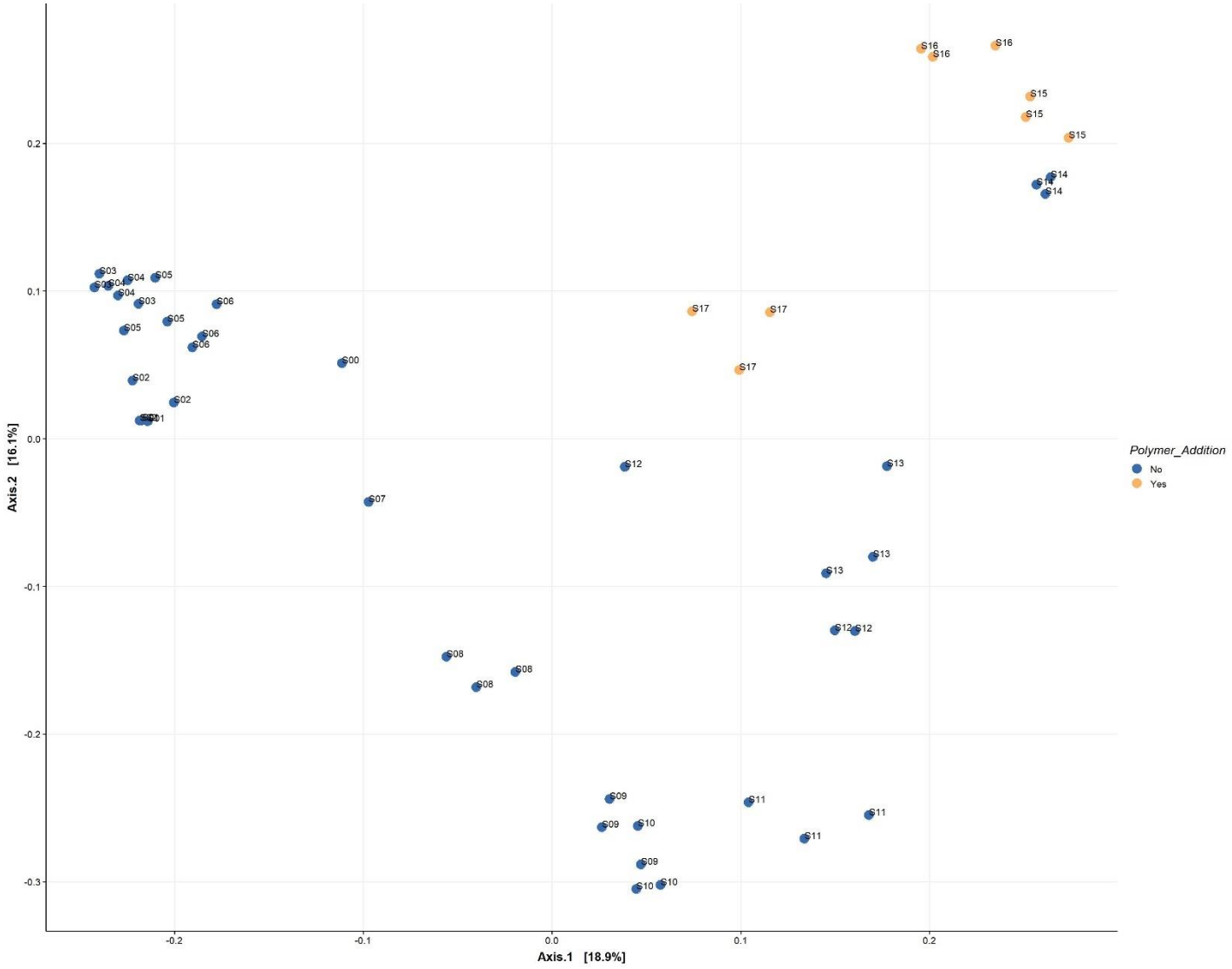
SI Figure F.27: Unweighted UniFrac beta diversity plot for GAC10 with pilot runtime identified.



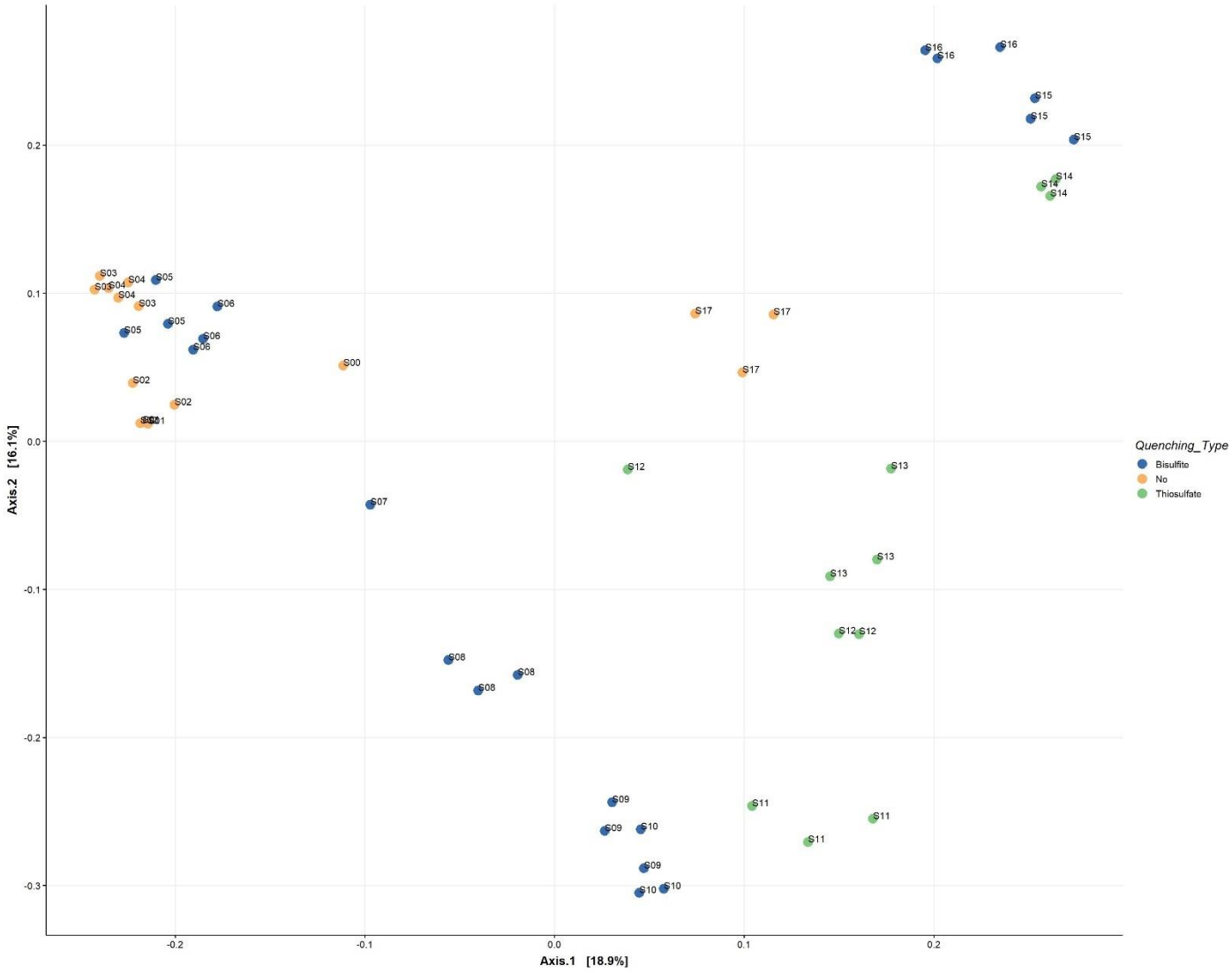
SI Figure F.28: Unweighted UniFrac beta diversity plot for GAC10 with temporal proximity to backwash identified.



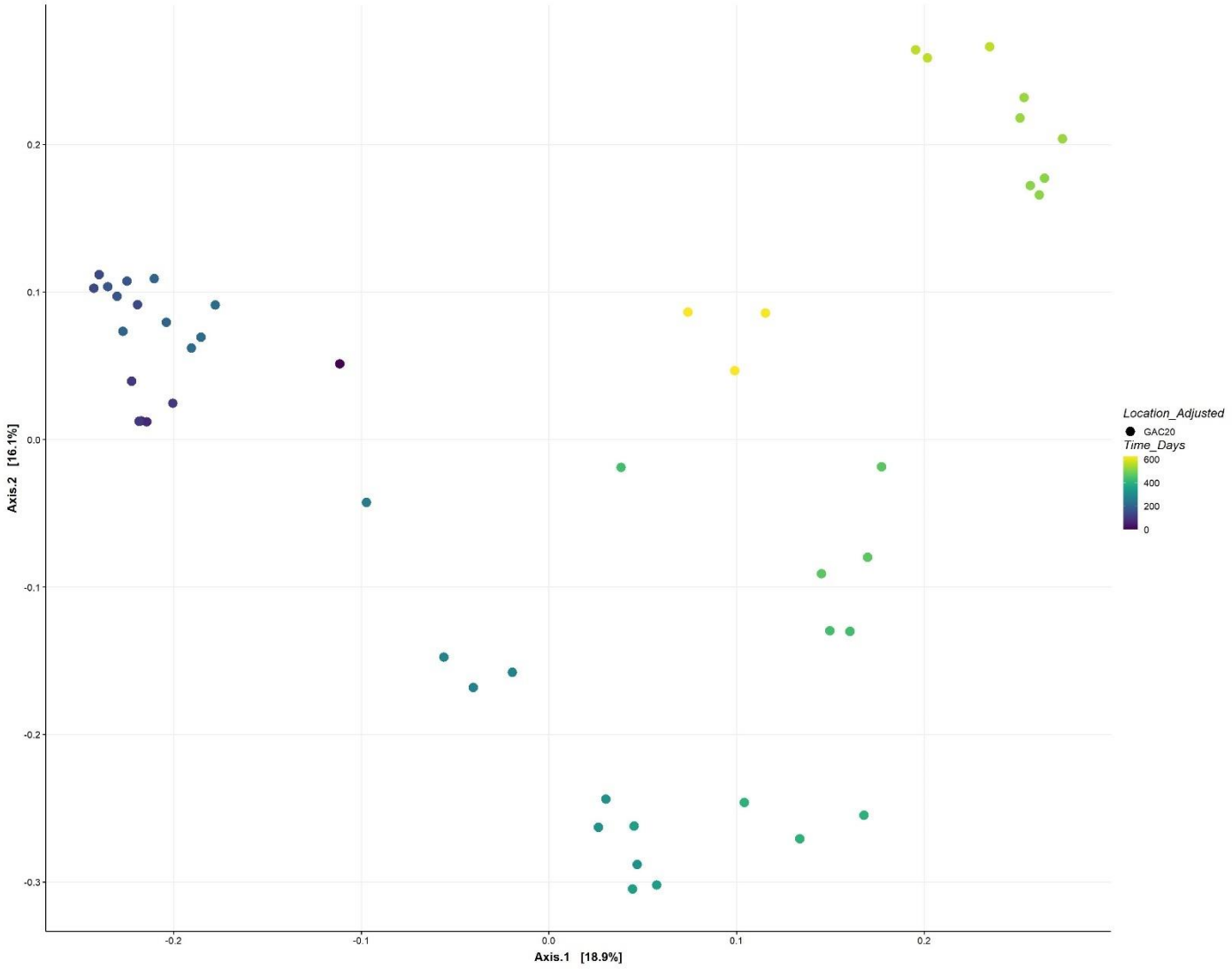
SI Figure F.29: Unweighted UniFrac beta diversity plot for GAC20 with adjusted O_3:TOC ratio identified.



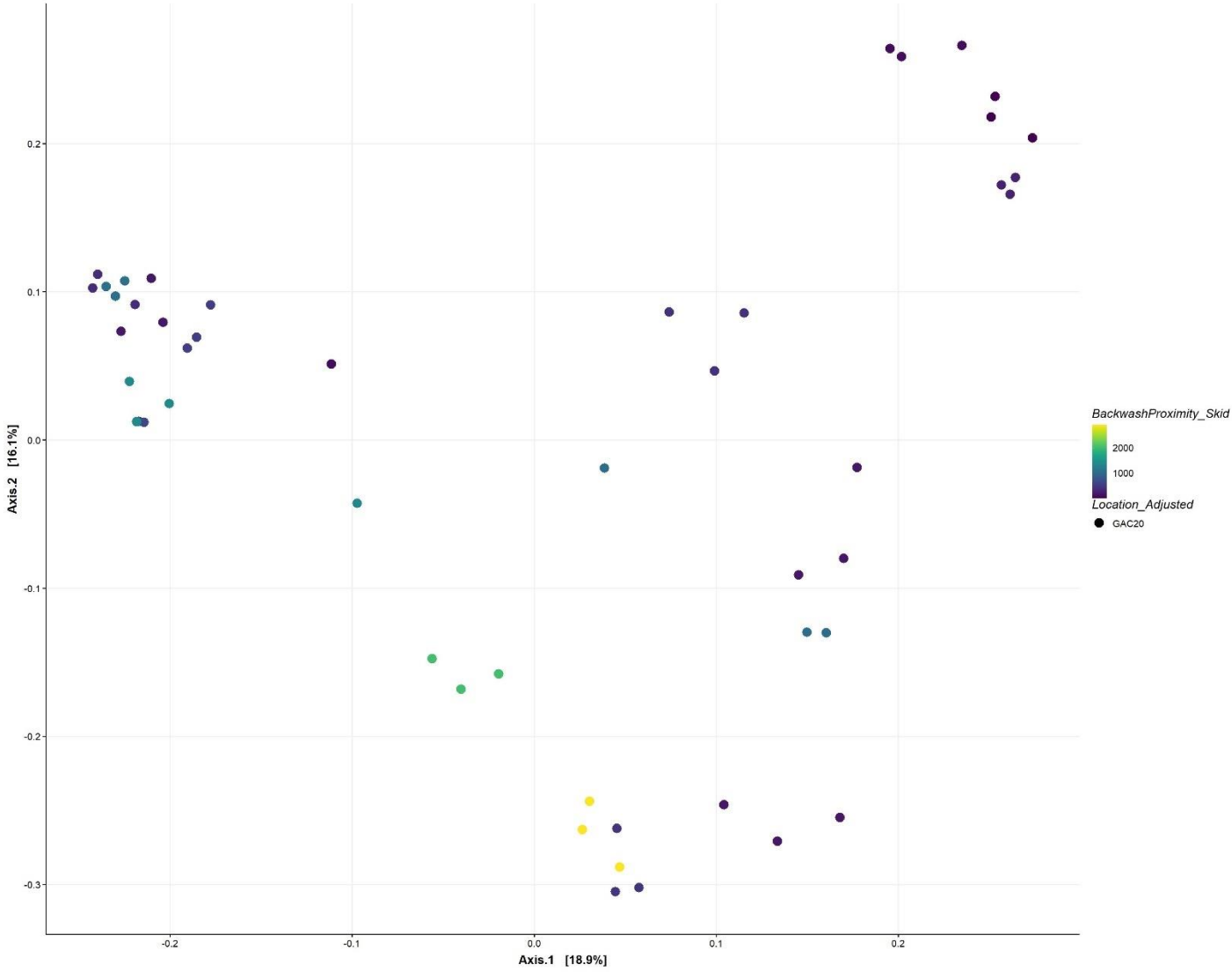
SI Figure F.30: Unweighted UniFrac beta diversity plot for GAC20 with polymer addition identified.



SI Figure F.31: Unweighted UniFrac beta diversity plot for GAC20 with quenching type identified.

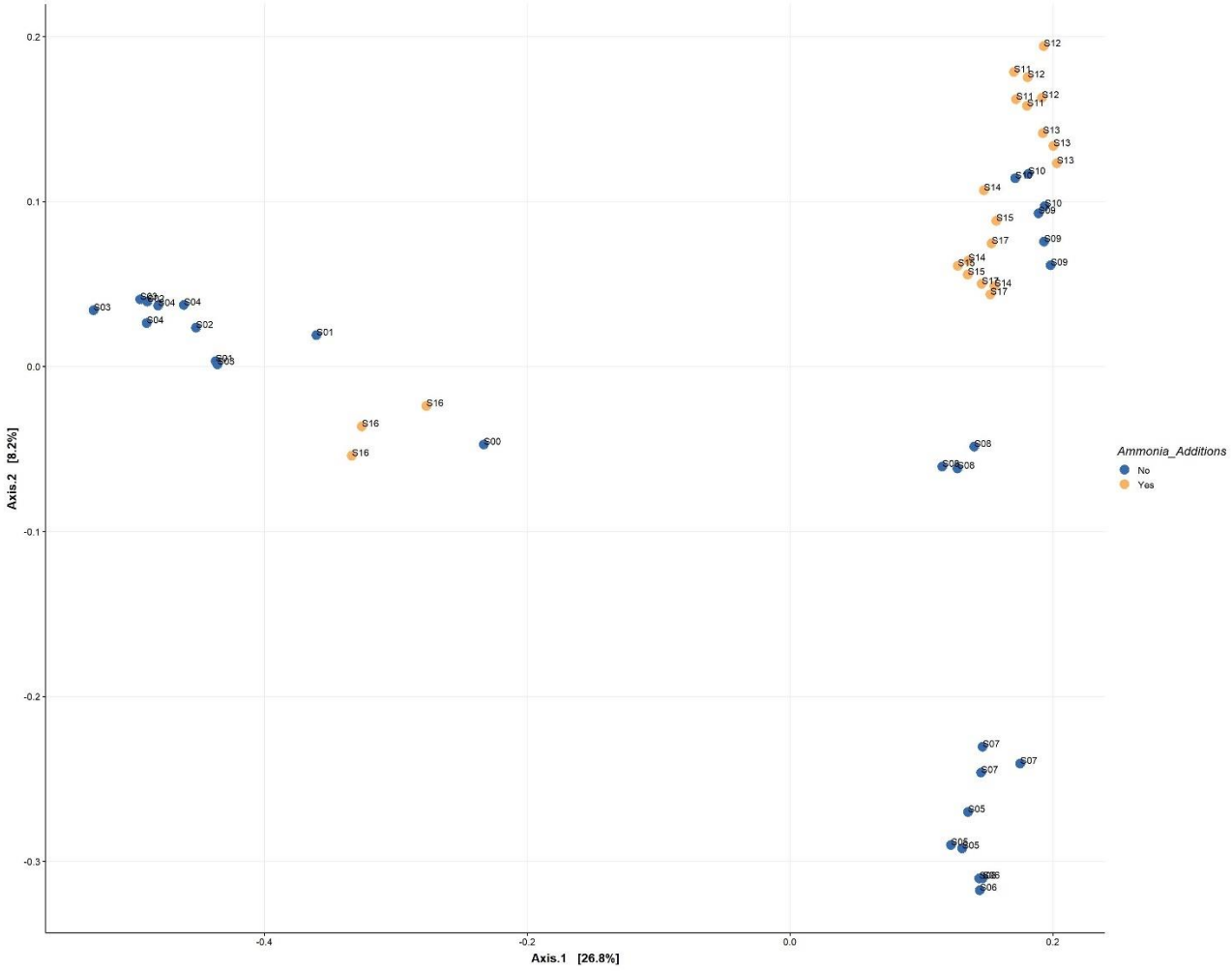


SI Figure F.32: Unweighted UniFrac beta diversity plot for GAC20 with pilot runtime identified.

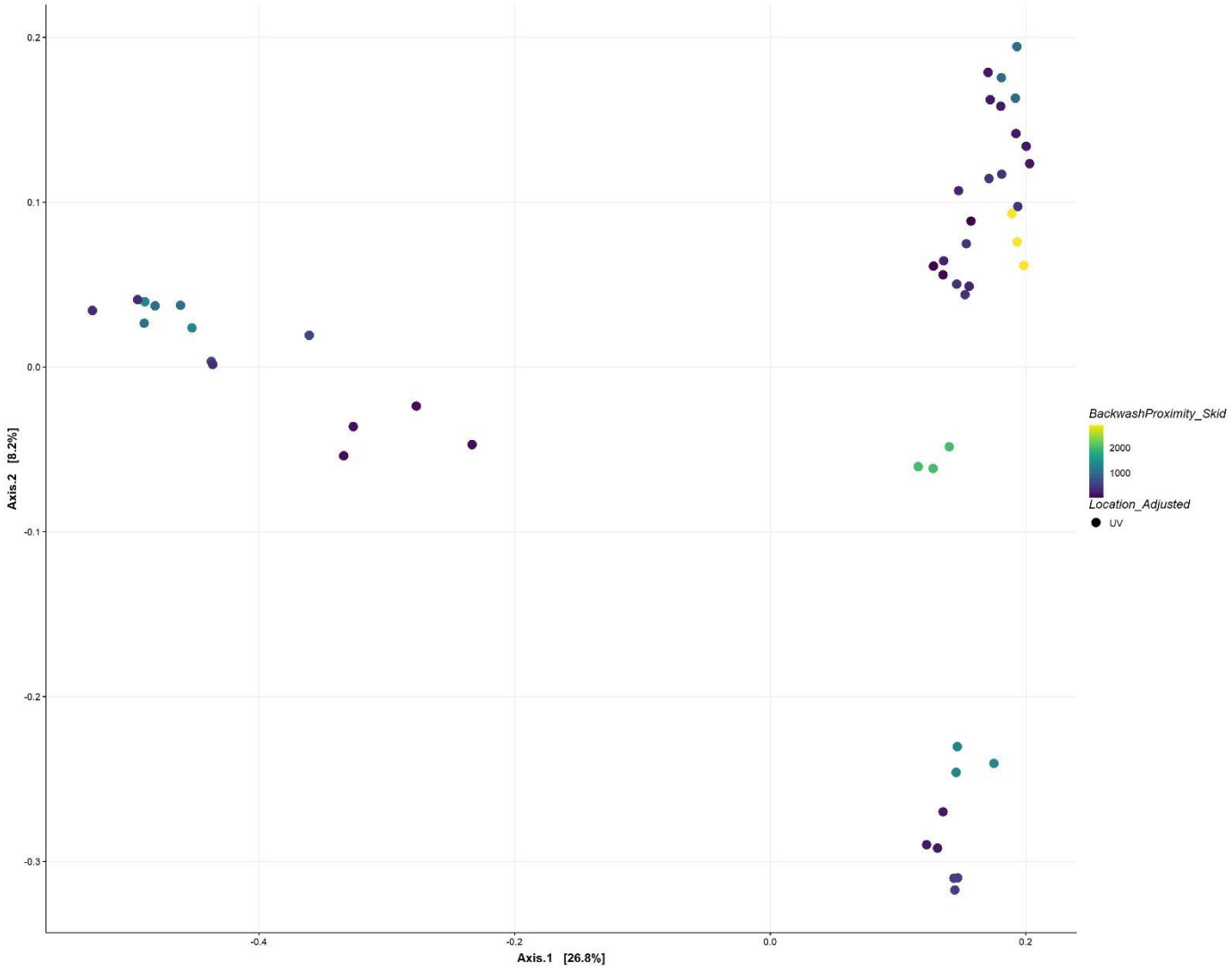


SI Figure F.33: Unweighted UniFrac beta diversity plot for GAC20 with temporal proximity to backwash identified.

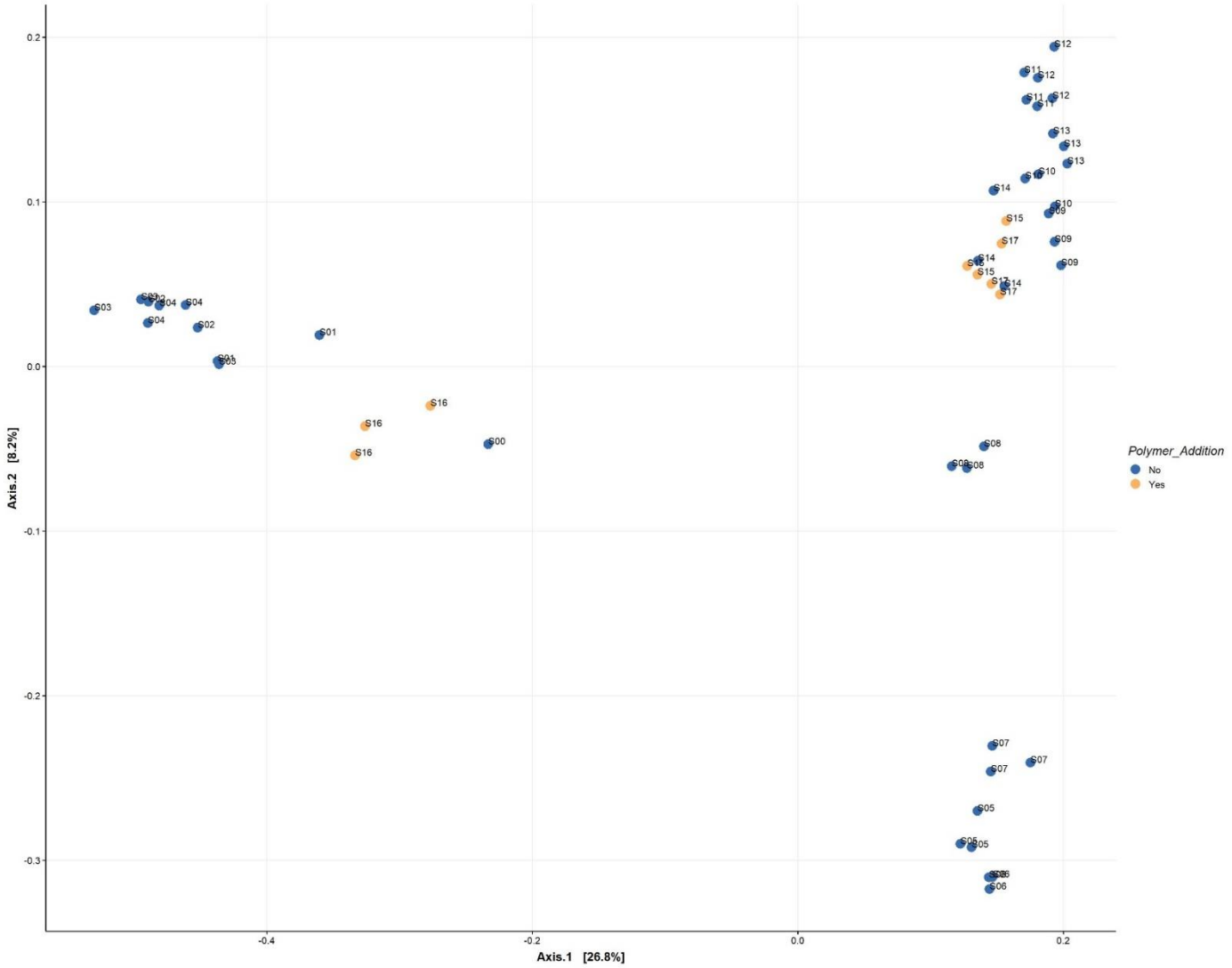
Subsection 5: Supplemental Ordination Plots: UV Irradiation



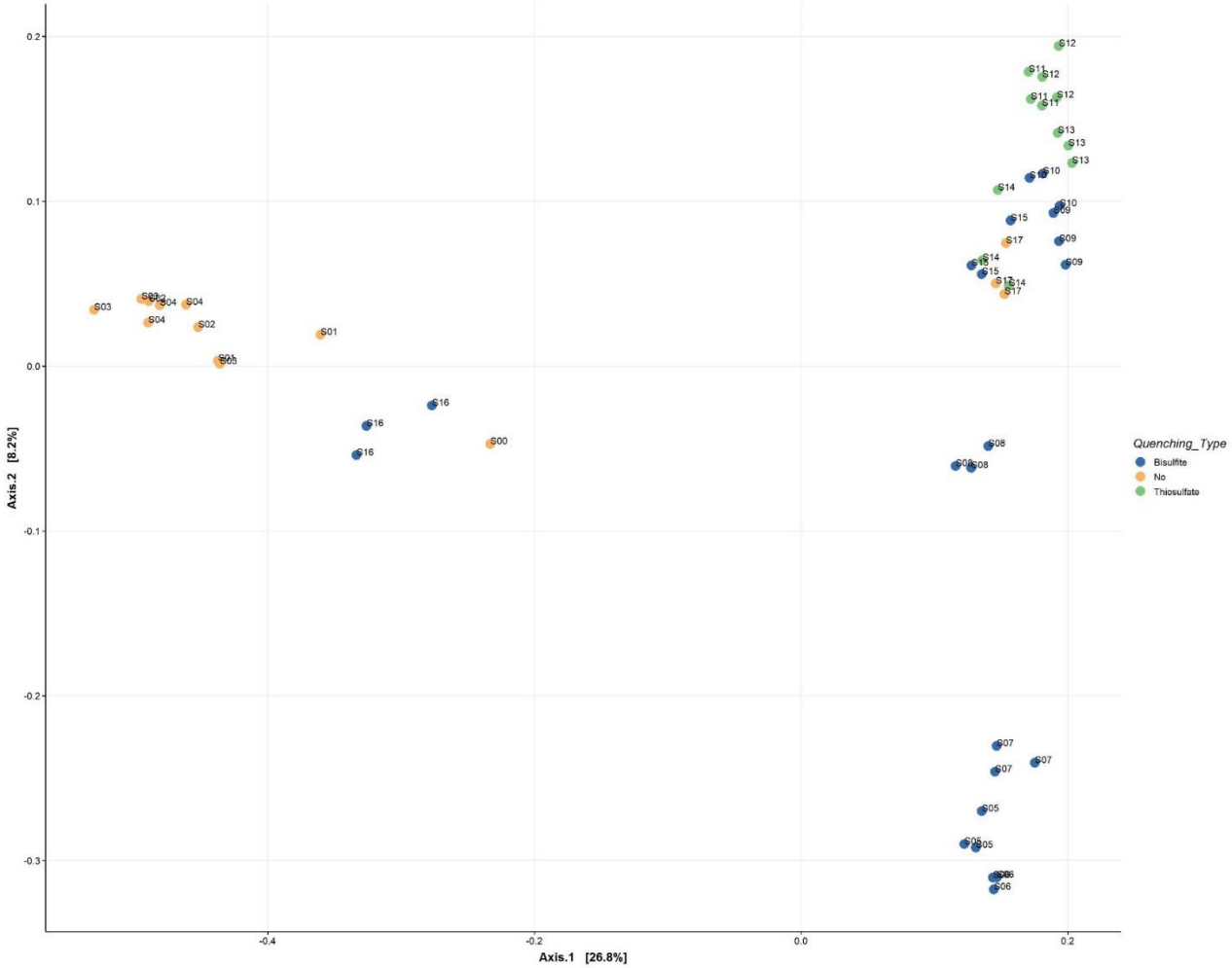
SI Figure F.34: Unweighted UniFrac beta diversity plot with supplemental nutrient additions identified.



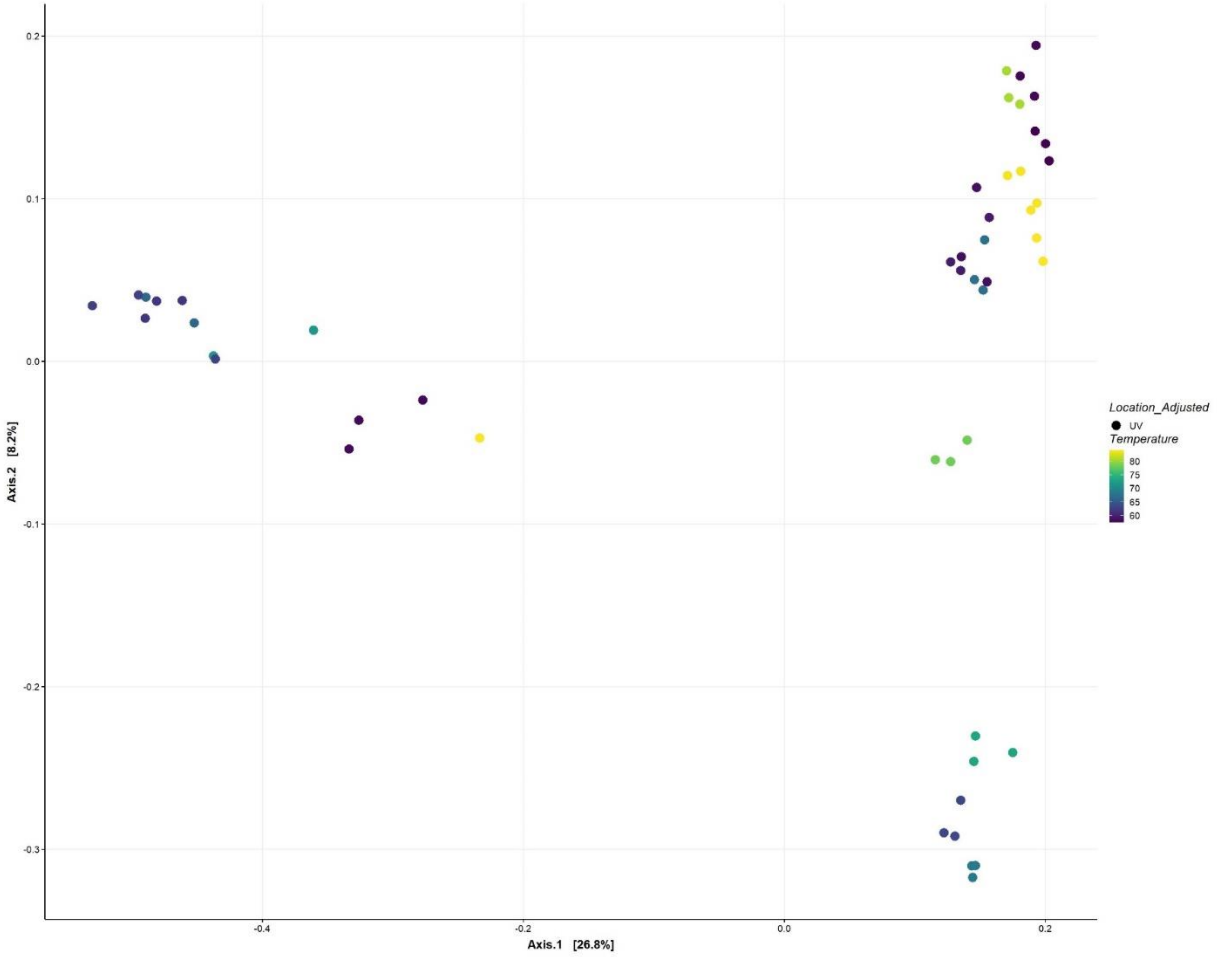
SI Figure F.35: Unweighted UniFrac beta diversity plot with temporal proximity to backwash identified.



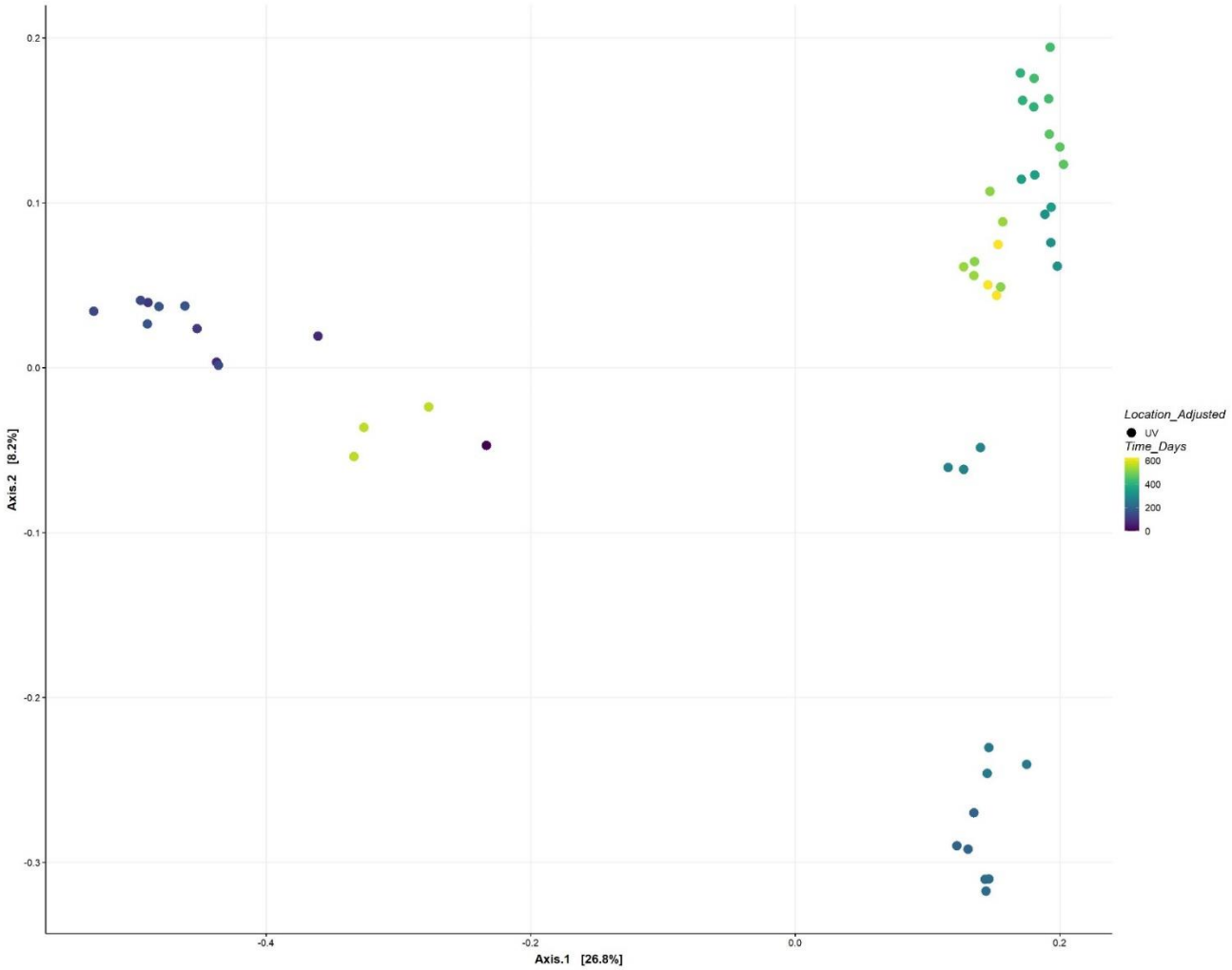
SI Figure F.36: Unweighted UniFrac beta diversity plot with polymer addition identified.



SI Figure F.37: Unweighted UniFrac beta diversity plot with quenching type identified.



SI Figure F.38: Unweighted UniFrac beta diversity plot with temperature identified.



SI Figure F.39: Unweighted UniFrac beta diversity plot with pilot runtime identified.

APPENDIX G: SUPPLEMENTAL INFORMATION FOR CHAPTER 6

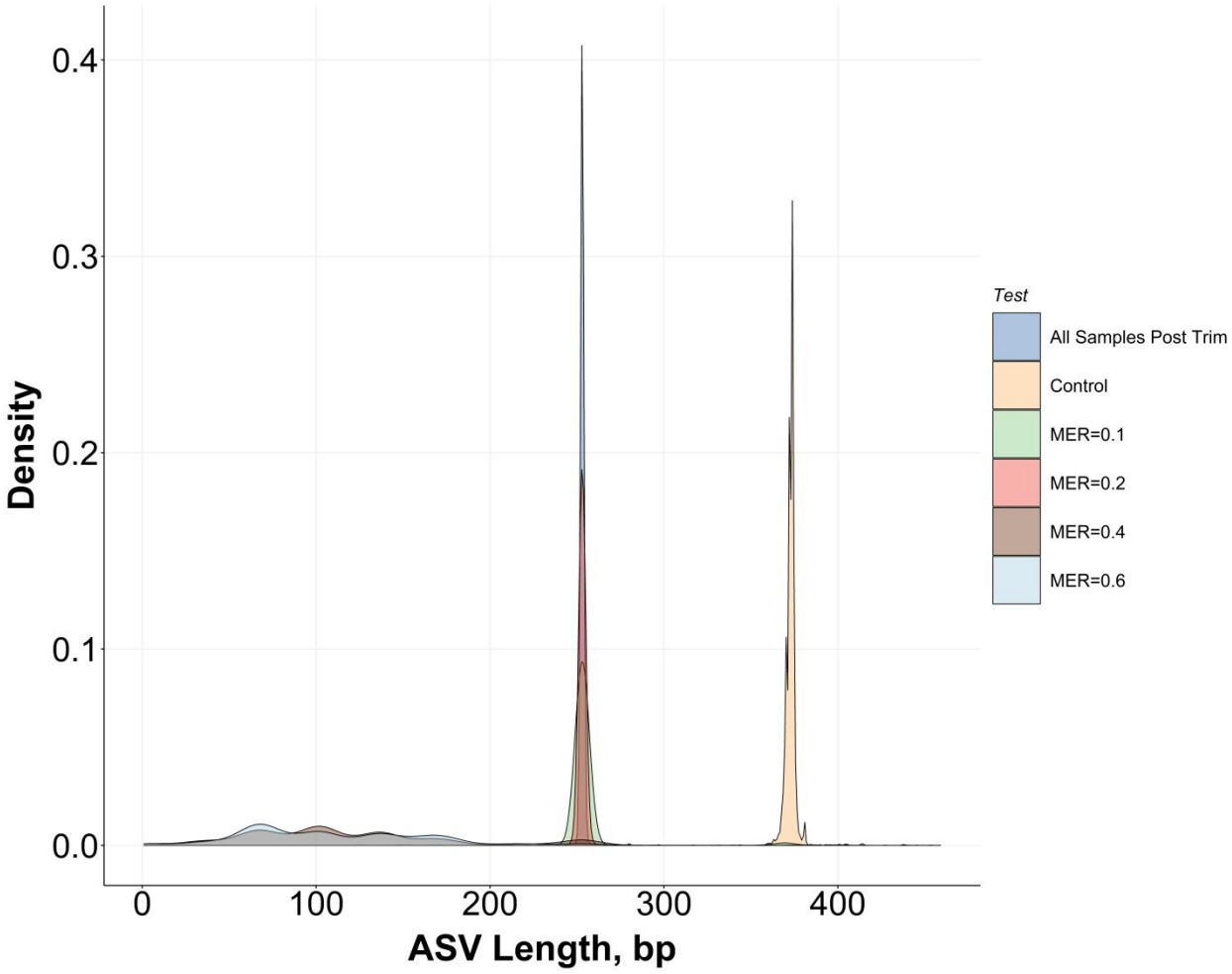
Subsection 1: Validation of Post-Sequencing Trimming

One of the first objectives of this work was to validate consolidated analysis of 16S rRNA gene amplicon sequences originating from different primers with an overlapped sequencing region. As NGS technologies continue to gain popularity and be applied throughout the water and wastewater industries it is inevitable that variation will be introduced throughout all stages of sequencing without an agreed upon standard. Variation present among downstream analysis is the least worrisome since results can always be reproduced in accordance with agreed upon methods as standards become adopted industry-wide. However, variability present prior to sequencing (e.g. within DNA extraction, library prep, and sequencing pipelines) have a much greater potential to introduce variation that cannot be accounted for following the completion of sequencing. Therefore, to increase the longevity and usability of sequencing results and allow for an increase in industry wide collaborative studies, there is a need to be able to develop and validate techniques to overcome minor differences in preparatory pipelines prior to sequencing. One such case relates to the variability associated with differing targeted regions of the 16S rRNA gene utilized in 16Sco rRNA gene amplicon sequencing.

Here sequencing results from two commonly targeted regions (V4 and V4-V5) (with an overlapped amplicon region) were successfully combined by trimming produced sequence variants down to an identified primer binding site using the cut-adapt tool. SI Figure A.1 graphically presents the distribution of base pairs (bps) lengths for each generated ASV, before and after multiple attempts at trimming. The control samples represent a lane of untrimmed sequences produced from the 515f-926r primer set targeting both the V4-V5 region, while the combined samples represent the ideal targeted range of ASV lengths when only using the V4 region and 515f-806r primer set. Similarly, SI Figure G.2 provides the distribution of removed bps, measured via the change in sequence length, during multiple trimming attempts. Trims around 120 bps were ideal as the 515f-926r primer set targets a region that is roughly 120 bps longer than the 515f-806r primer set. However, it is worth noting that various organisms contain nucleotide variation throughout the variable region(s) of the 16S rRNA gene indicating that all sequences will not have exactly the same ASV length nor the same change in length. Instead a reasonably tight distribution about the 120-bp trim was expected. Changes in ASV length much greater than 120 bps or much lower than 120 bps indicated too lenient or strict trimming parameters, respectively.

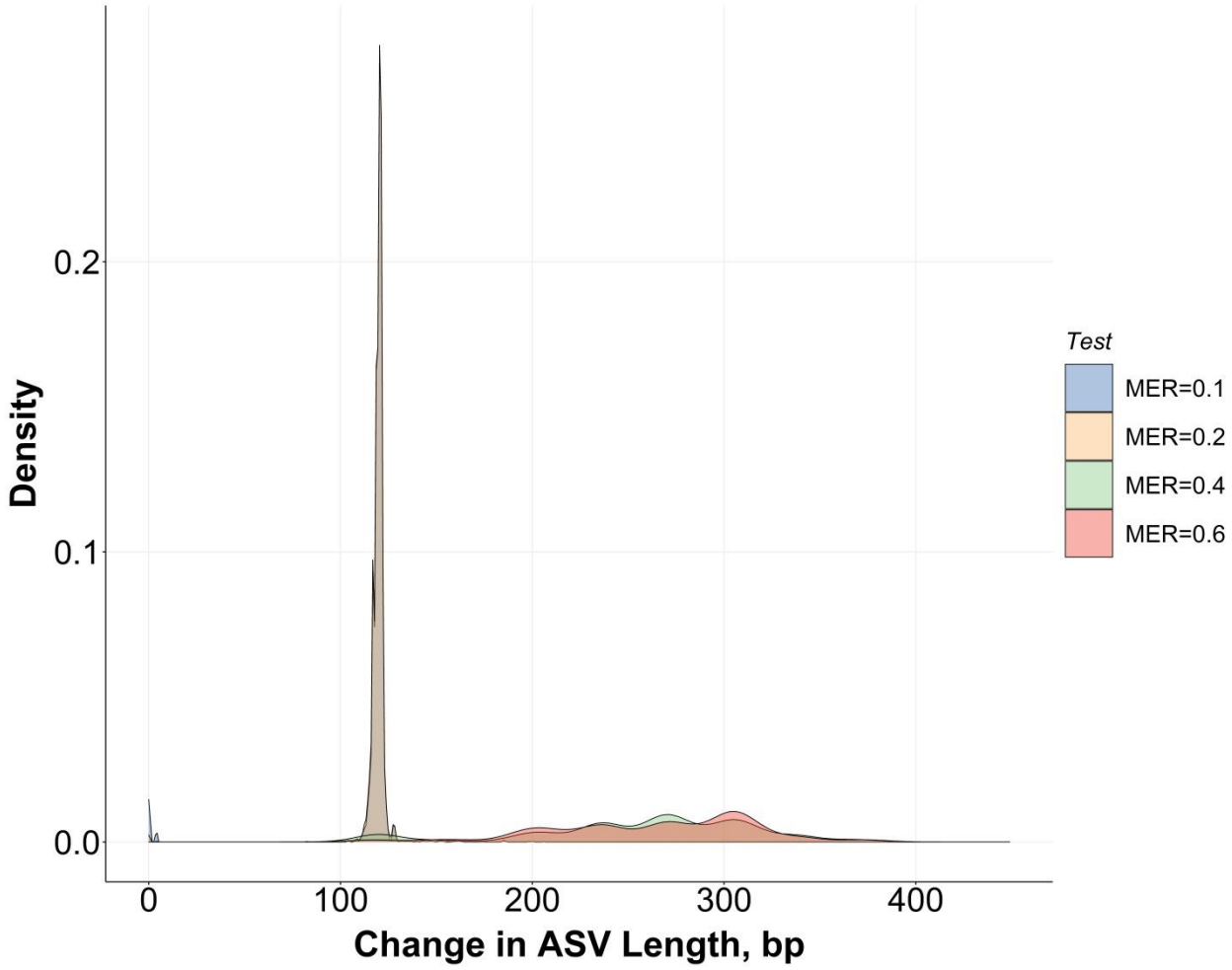
With this context understood, SI Figure A.1 and SI Figure G.2 support the ability for the Cutadapt tool to be utilized ex post facto, with optimized trimming parameters. Trimming was effective at retroactively reducing 515f-926r generated ASVs down to the 806r binding site, consolidating all generated sequences to the same V4 region of the 16S rRNA gene. Changes in the Cutadapt MER were seen to have the greatest impact on the effectiveness of trimming. Increased MERs lead to decreased binding site specificity and excessive trimming. Conversely, a lower max error rate increased primer site specifically and lead to more accurate trims with resultant ASVs falling within the desirable trimming ranges. However, lower MERs also lead to a higher percentage of untrimmed sequences. Therefore, a tiered trimming approach was found to be most effective, where sequences were initially trimmed with a low max error rate

(MER=0.1) and then untrimmed sequences were re-trimmed with a higher max error rate (MER=0.2). The resultant sequences from the tiered approach proved to be the only resultant sequences tested that did not have sequence length that were statistically different from the targeted lengths (ANOVA, p-value = 0.987) with all other trimming approaches (MER = 0.1, 0.4, 0.6) generating statistically different length distributions (ANOVA, p-value < 0.05), SI Table G.1. Further, all tested trimming parameters generated statistically different change in length distributions (ANOVA, p-value < 0.05) indicating that the Cutadapt tool's MER parameter was an effective means of controlling trimming site specificity, SI Table G.2. The tiered approach with final MER=0.2 resulted in 99.32% of all sequencing being trimmed correctly, 0.38% remaining untrimmed, and 0.30% being poorly trimmed. This compared favorably to 96.98% correctly trimmed, 2.32% untrimmed, and 0.70% poorly trimmed for the default Cutadapt parameters, SI Table G.3 and SI Table G.4. Trimming validation asserts that post sequencing trimming is a reasonable and effective approach to consolidating 16S rRNA gene amplicon sequences originating from different primers with an overlapped sequencing region into a single region of interest, prior to downstream analysis.



SI Figure G.1: Density plot of 16S rRNA gene amplicon ASV lengths after multiple Cutadapt pipelines trimmed ASVs generated from the 515f/926r down to the 515f/806r binding site. Here the control density represents untrimmed 515f/826r ASV lengths, all

samples post trim represent the model 515f/806r reference density, MER=0.1 represents the default Cutadapt parameters, and MER=0.2, MER=0.4, and MER=0.6 represents a tiered trimming approach with the secondary trim MER set to their respective values after an initial trim at the default parameters.



SI Figure G.2: Density plot of the removed base pairs for each Cutadapt pipeline when trimming 16S rRNA gene amplicon ASV lengths from the 515f/926r primer set down to the 515f/806r binding site. MER stand for max error rate and is a parameter utilized by the cutadapt tool. Default MER is set to 0.1 while experimental testing targeted supplementary MERs of 0.2, 0.4, and 0.6, as outlined in the methods.

Subsection 2: Supplemental Tables and Figures

SI Table G.1: ANOVA results for the comparison of ASV length distributions post trimming.

ANOVA Comparison	P-value
Control-All Samples Post Trim	0.000

MER=0.1-All Samples Post Trim	0.000
MER=0.2-All Samples Post Trim	0.988
MER=0.4-All Samples Post Trim	0.000
MER=0.6-All Samples Post Trim	0.000
MER=0.1-Control	0.000
MER=0.2-Control	0.000
MER=0.4-Control	0.000
MER=0.6-Control	0.000
MER=0.2-MER=0.1	0.000
MER=0.4-MER=0.1	0.000
MER=0.6-MER=0.1	0.000
MER=0.4-MER=0.2	0.000
MER=0.6-MER=0.2	0.000
MER=0.6-MER=0.4	0.000

SI Table G.2: ANOVA results for the change in ASV length post trimming.

ANOVA Comparison	P-value
MER=0.2-MER=0.1	0.000
MER=0.4-MER=0.1	0.000
MER=0.6-MER=0.1	0.000
MER=0.4-MER=0.2	0.000
MER=0.6-MER=0.2	0.000
MER=0.6-MER=0.4	0.000

SI Table G.3: Post trimming results based on ASV length.

ASV Length, bp	Untrimmed Control, n=9,315	MER 0.1, n=9,315	MER 0.1 + MER 0.2, n=9,315	MER 0.1 + MER 0.4, n=9,315	MER 0.1 + MER 0.6, n=9,315	All Samples Post Trim, n=71,119
<225	0.00%	0.01%	0.19%	90.89%	96.55%	0.12%
226-275	0.00%	97.16%	99.32%	9.02%	3.39%	99.53%
276-349	0.44%	0.49%	0.20%	0.06%	0.03%	0.26%
>350	99.56%	2.34%	0.28%	0.03%	0.02%	0.09%
Total	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**

Note: MER stands for Max Error Rate and is a parameter within the cutadapt tool

SI Table G.4: Post trimming results based on the change in ASV lengths.

Amount Trimmed, bp	MER 0.1, n=9,315	MER 0.1 + MER 0.2, n=9,315	MER 0.1 + MER 0.4, n=9,315	MER 0.1 + MER 0.6, n=9,315
0, Untrimmed	2.32%	0.38%	0.00%	0.00%
100-140, Ideal	96.67%	98.85%	7.86%	2.33%
80-100 & 140-160, Transitional	0.31%	0.47%	1.81%	1.92%
1-79 & 161+, Poor	0.70%	0.30%	90.33%	95.75%
Total	**100.00%**	**100.00%**	**100.00%**	**100.00%**

Note: MER stands for Max Error Rate and is a parameter within the cutadapt tool

SI Table G.5: Summary of Climate metadata based on intended water use.

Köppen Climate Classification	Description	Potable	Non-Potable Reuse	Reuse	Totals
Bsh	Mid-latitude steppe and desert climate	9	19	0	28
Bwk	Cold desert climate	3	0	12	15
Cfa	Humid subtropical climate	78	107	15	200
Csa	Hot-summer Mediterranean climate	6	0	8	14
Csb	Warm-summer Mediterranean climate	39	51	18	108
Dfa	Hot-summer humid continental climate	20	0	0	20
Dfb	Warm-summer humid continental climate	40	0	0	40
	Total	195	177	53	425

SI Table G.6. Summary of disinfection processes based on intended water use.

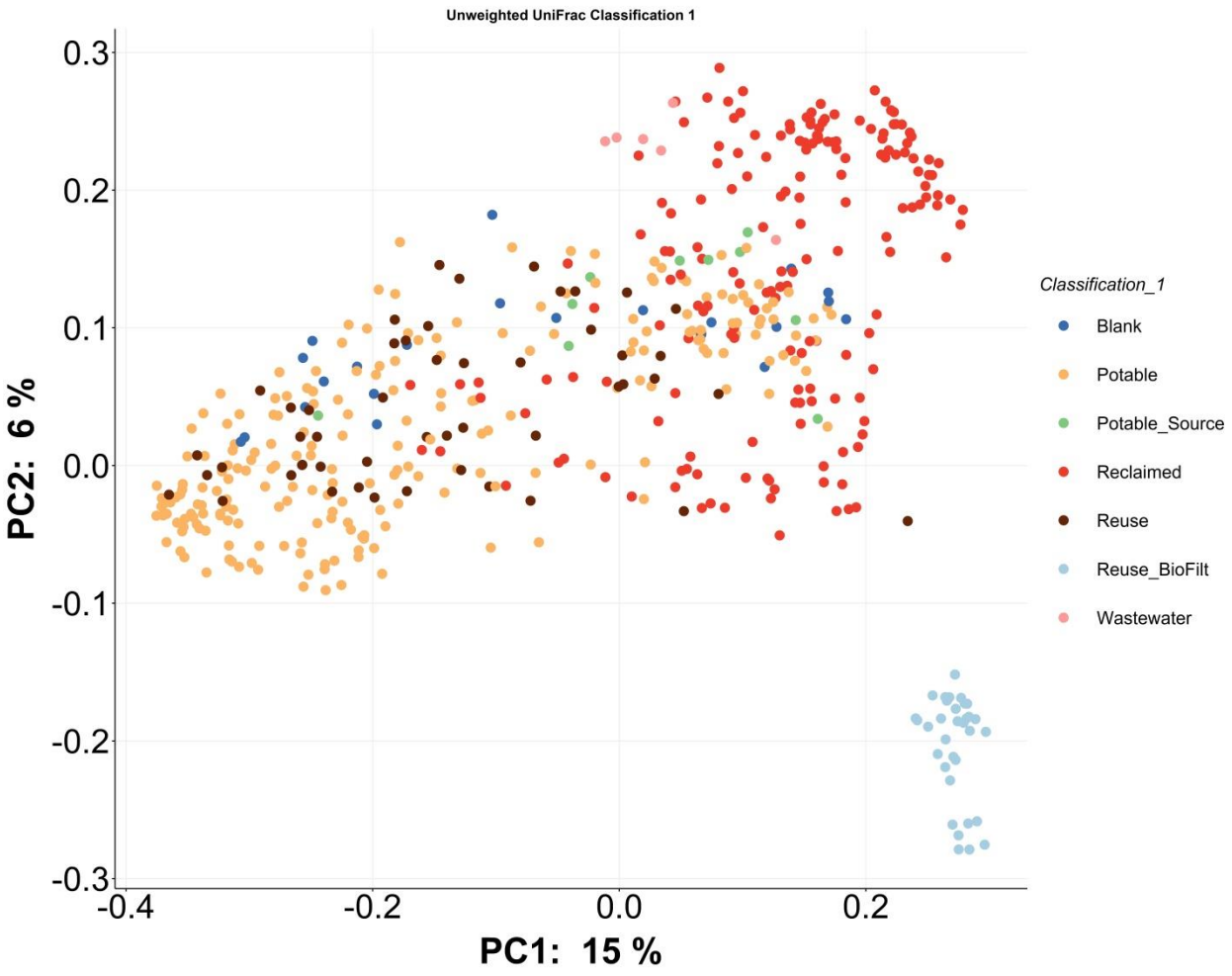
Disinfection Treatments	Potable	Non-Potable Reuse	Potable Reuse	Totals
Chloramine	41	19	0	60
Chlorine	90	0	6	96
Chlorine Chloramine	0	98	0	98
Chlorine ChlorineDioxide	18	0	0	18
Ozone Chloramine Industry	0	0	3	3
Ozone Chlorine	3	0	9	12
Ozone Chlorine Chloramine	3	0	0	3
Ozone UV Chloramine	0	0	15	15
Ozone UV Chlorine	0	0	2	2
Ozone UV Chlorine Chloramine	0	0	3	3
UV	0	0	4	4
UV Chloramine	20	4	2	26
UV Chlorine	20	0	6	26
UV Chlorine Pasteurization	0	0	3	3

SI Table G.7. Summary of disinfectant residuals based on intended water use

Disinfection Residual	Potable	Non-Potable Reuse	Potable Reuse	
Chloramine	70	121	26	217
Chlorine	125	0	12	137

SI Table G.8: Summary of POC vs POU samples based on intended water use.

Sample Location	Potable	Non-Potable Reuse	Potable Reuse
Point of Compliance, POC (WTP)	63	82	27
Point of Use, POU (DS)	132	95	26



SI Figure G.3: Unweighted UniFrac beta diversity plot for all bulk water samples regardless of distinctions between POC and POU, classified by intended water use with reuse biofiltration reference community.

SI Table G.9: ANOSIM p-value results for comparison made via water use and sample location distinctions.

P-value	POC and POU Samples				POC Samples				POU Samples			
	Combined	Potable	Non-Potable Reuse	Reuse	Combined	Potable	Non-Potable Reuse	Reuse	Combined	Potable	Non-Potable Reuse	Reuse
Water Use	0.001	-	-	-	0.001	-	-	-	0.001	-	-	-
Water Use, High Resolution	0.001	0.004	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.005	-	-
Climate	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.111
Disinfection Residual	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	-	0.262
Region	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	0.001	0.001	0.001	-

Location	0.001	0.084	0.001	0.001	-	-	-	-	-	-	-	-

Note: P-values < 0.001 were entered as 0.001

SI Table G.10: ANOSIM r-statistic results from comparisons made via water use and sample location distinctions.

R-Statistic	POC and POU Samples				POC Samples				POU Samples			
	Combined	Potable	Non-Potable Reuse	Reuse	Combined	Potable	Non-Potable Reuse	Reuse	Combined	Potable	Non-Potable Reuse	Reuse
Water Use	0.450	-	-	-	0.746	-	-	-	0.339	-	-	-
Water Use, High Resolution	0.419	0.083	0.386	0.205	0.767	0.360	0.795	0.583	0.346	0.155	-	
Climate	0.062	0.162	0.147	0.283	0.161	0.603	0.436	0.496	0.155	0.184	0.384	0.088
Disinfection Residual	0.281	0.253	0.338	0.234	0.509	0.345	0.795	0.486	0.200	0.264	-	0.028
Region	0.174	0.415	0.176	0.265	0.324	0.740	0.434	0.583	0.227	0.349	0.520	-
Location	0.071	0.038	0.356	0.164	-	-	-	-	-	-	-	-

Note: P-values < 0.001 were entered as 0.001

SI Table G.11: ANOSIM results for samples intended for potable use.

Potable Water Use	Combined		POC		POU	
	P-value	R-stat	P-value	R-stat	P-value	R-stat
Water Use, High Resolution	0.004	0.083	0.001	0.360	0.005	0.155
Climate	0.001	0.162	0.001	0.603	0.001	0.184
Disinfection Residual	0.001	0.253	0.001	0.345	0.001	0.264
Region	0.001	0.415	0.001	0.740	0.001	0.349
Location	0.084	0.038	-	-	-	-
Source Water	0.001	0.600	0.001	0.659	0.001	0.507
Treatment, Potable	0.001	0.535	0.001	0.563	0.001	0.531
Disinfection, Potable	0.001	0.324	0.001	0.355	0.001	0.364

Note: P-values < 0.001 were entered as 0.001

SI Table G.12: ANOSIM results for samples intended for non-potable reuse.

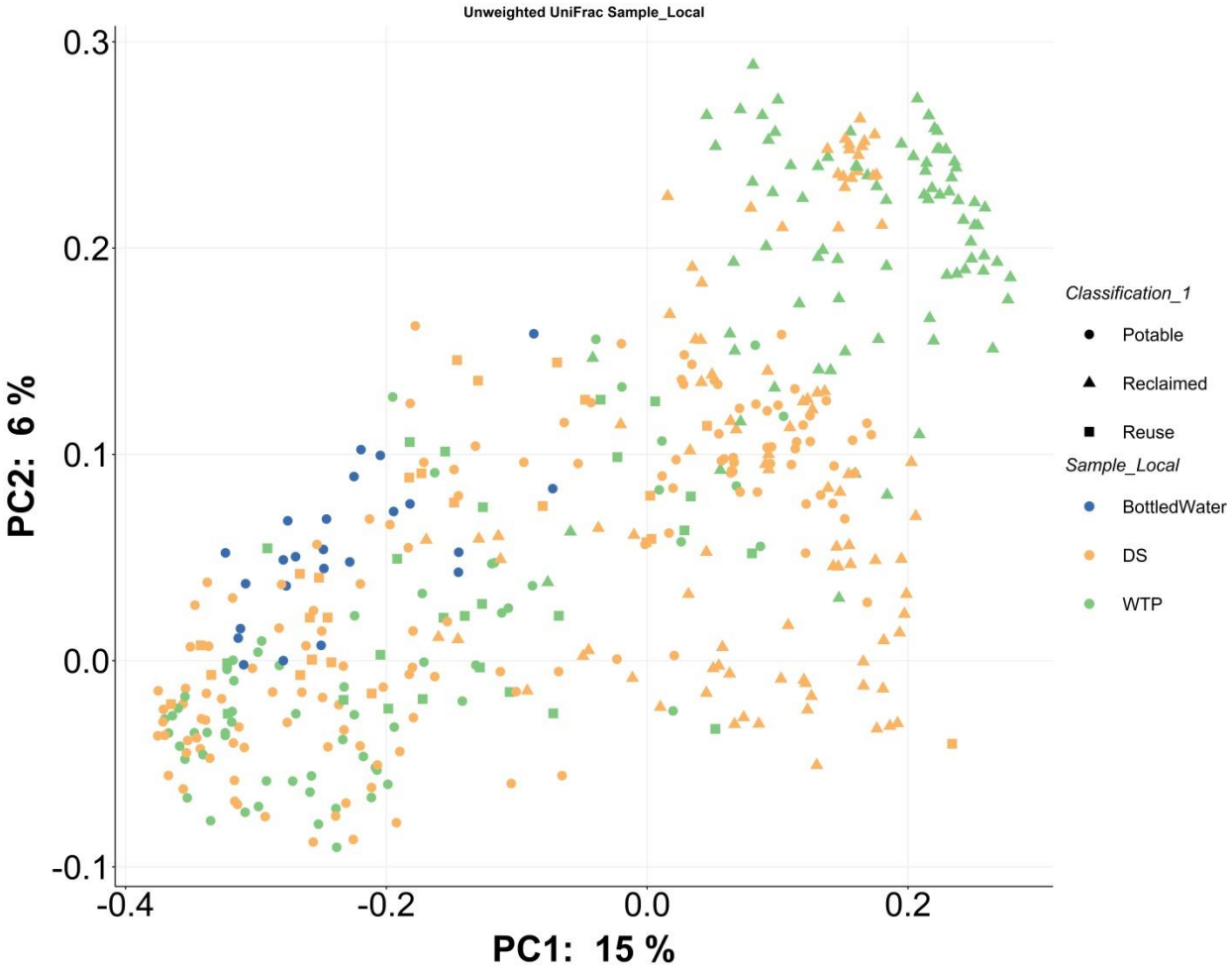
Non-potable Reuse	Combined		POC		POU	
	P-value	R-stat	P-value	R-stat	P-value	R-stat
Water Use, High Resolution	0.001	0.386	0.001	0.795	-	-
Climate	0.001	0.147	0.001	0.436	0.001	0.384
Disinfection Residual	0.001	0.338	0.001	0.795	-	-
Region	0.001	0.176	0.001	0.434	0.001	0.520
Location	0.001	0.356	-	-	-	-
Potable Reuse or Non-Potable Reuse, Disinfection	0.001	0.424	0.001	0.841	0.001	0.171
Potable Reuse or Non-Potable Reuse, Treatment	0.001	0.794	0.001	0.910	0.001	0.692

Note: P-values < 0.001 were entered as 0.001

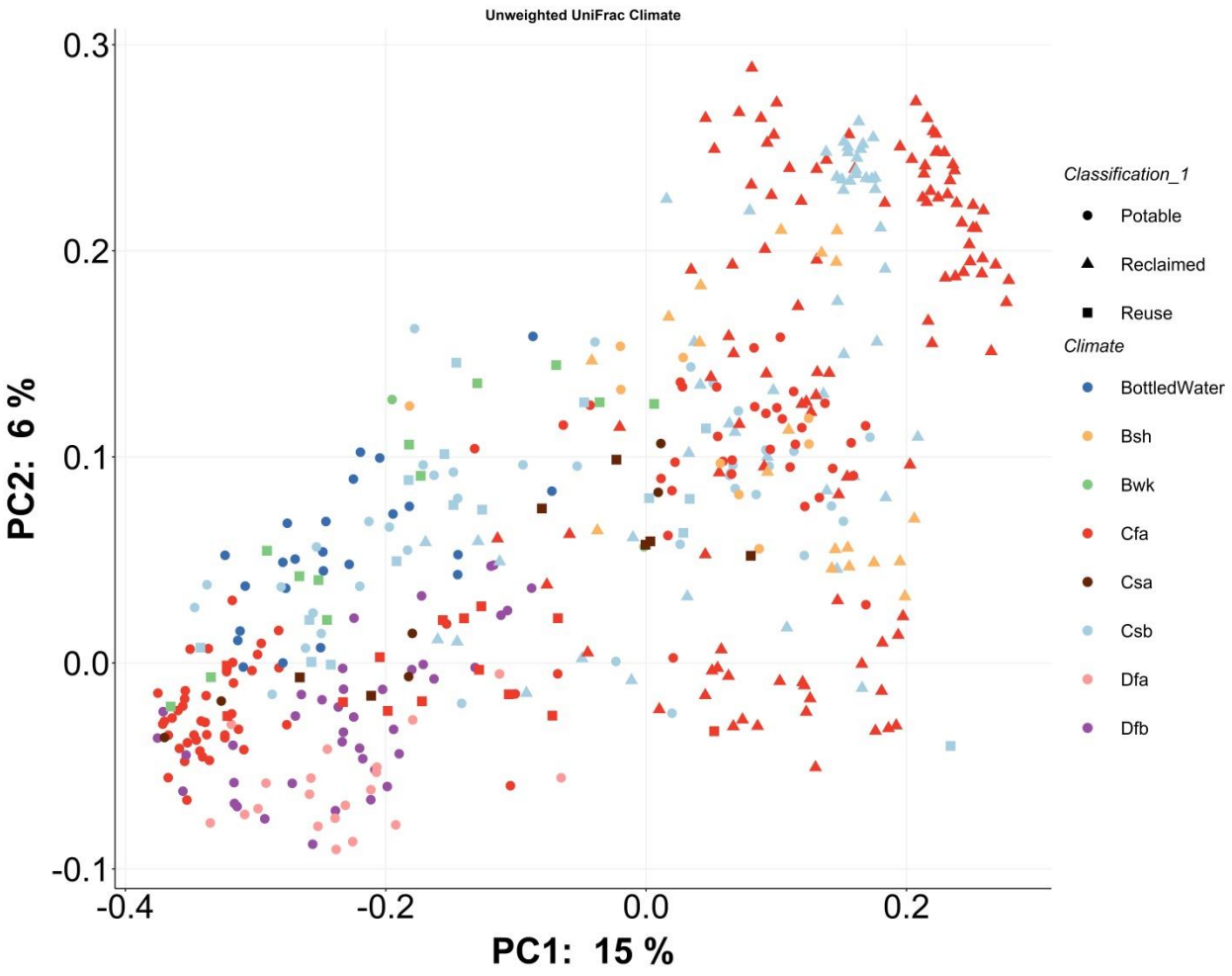
SI Table G.13: ANOSIM results for samples intended for potable reuse.

Potable Reuse	Combined		POC		POU	
	P-value	R-stat	P-value	R-stat	P-value	R-stat
Water Use, High Resolution	0.002	0.205	0.001	0.583	-	-
Climate	0.001	0.283	0.001	0.496	0.111	0.088
Disinfection Residual	0.001	0.234	0.001	0.486	0.262	0.028
Region	0.003	0.265	0.001	0.583	-	-
Location	0.001	0.164	-	-	-	-
Post Blending, Treatment	0.001	0.346	0.001	0.455	0.049	0.160
Blending Ratio	0.001	0.216	0.001	0.438	0.228	0.044
Potable Reuse or Non-Potable Reuse, Disinfection	0.001	0.325	0.001	0.461	0.085	0.135
Potable Reuse or Non-Potable Reuse, Treatment	0.001	0.325	0.002	0.461	0.075	0.135

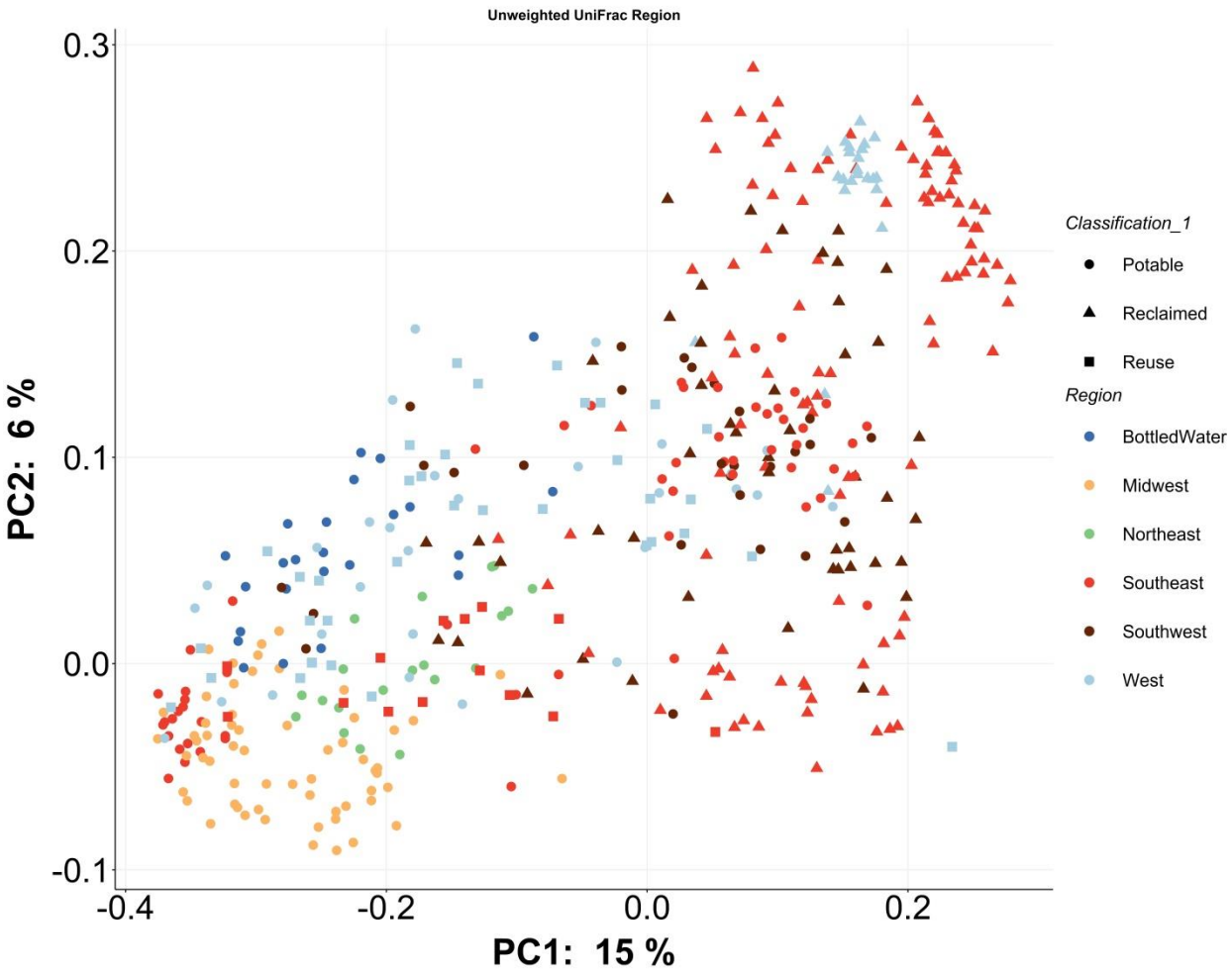
Note: P-values < 0.001 were entered as 0.001



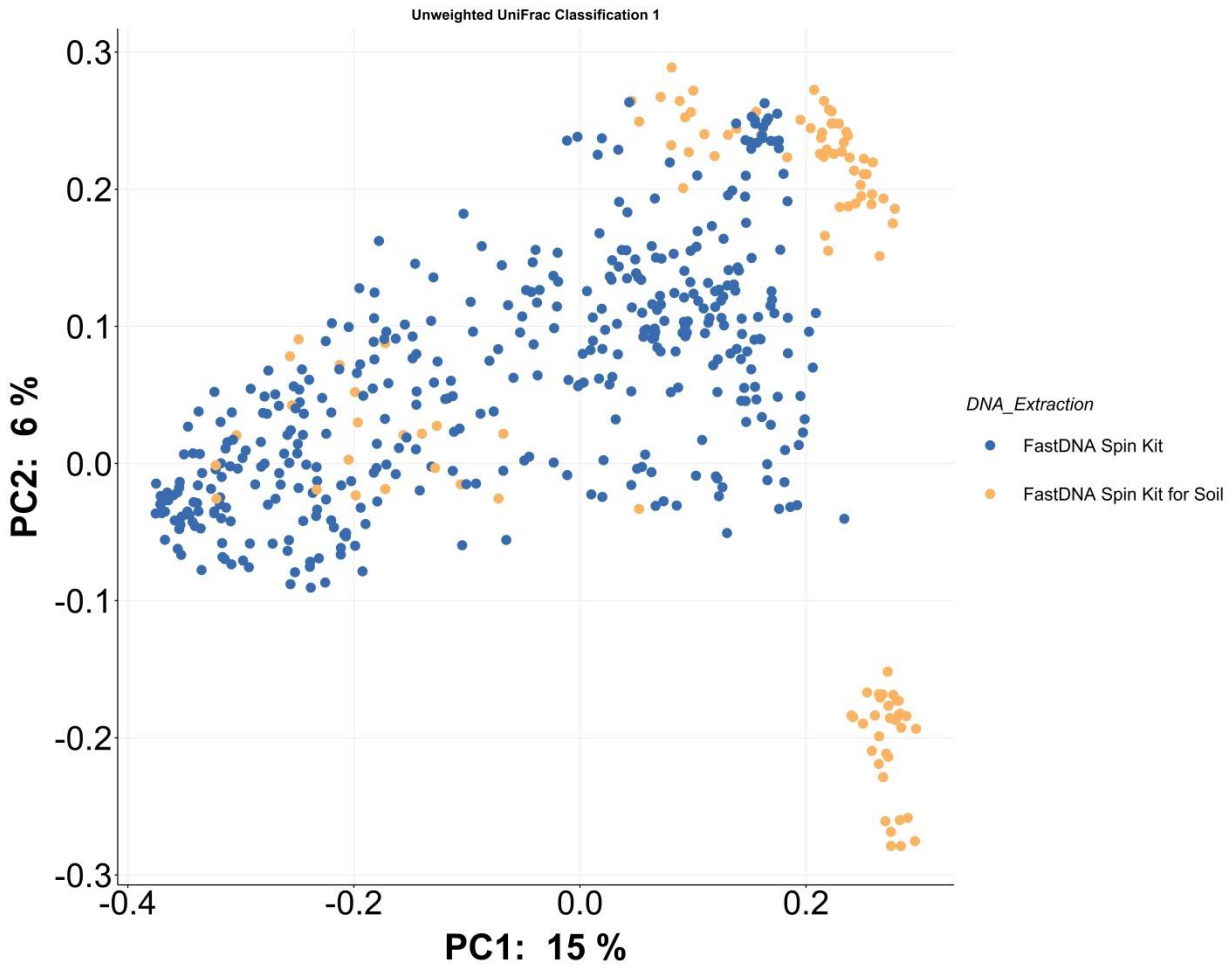
SI Figure G.4: Unweighted UniFrac beta diversity plot for all bulk water samples of interest, classified by intended water use and sample location related to the POC or POU.



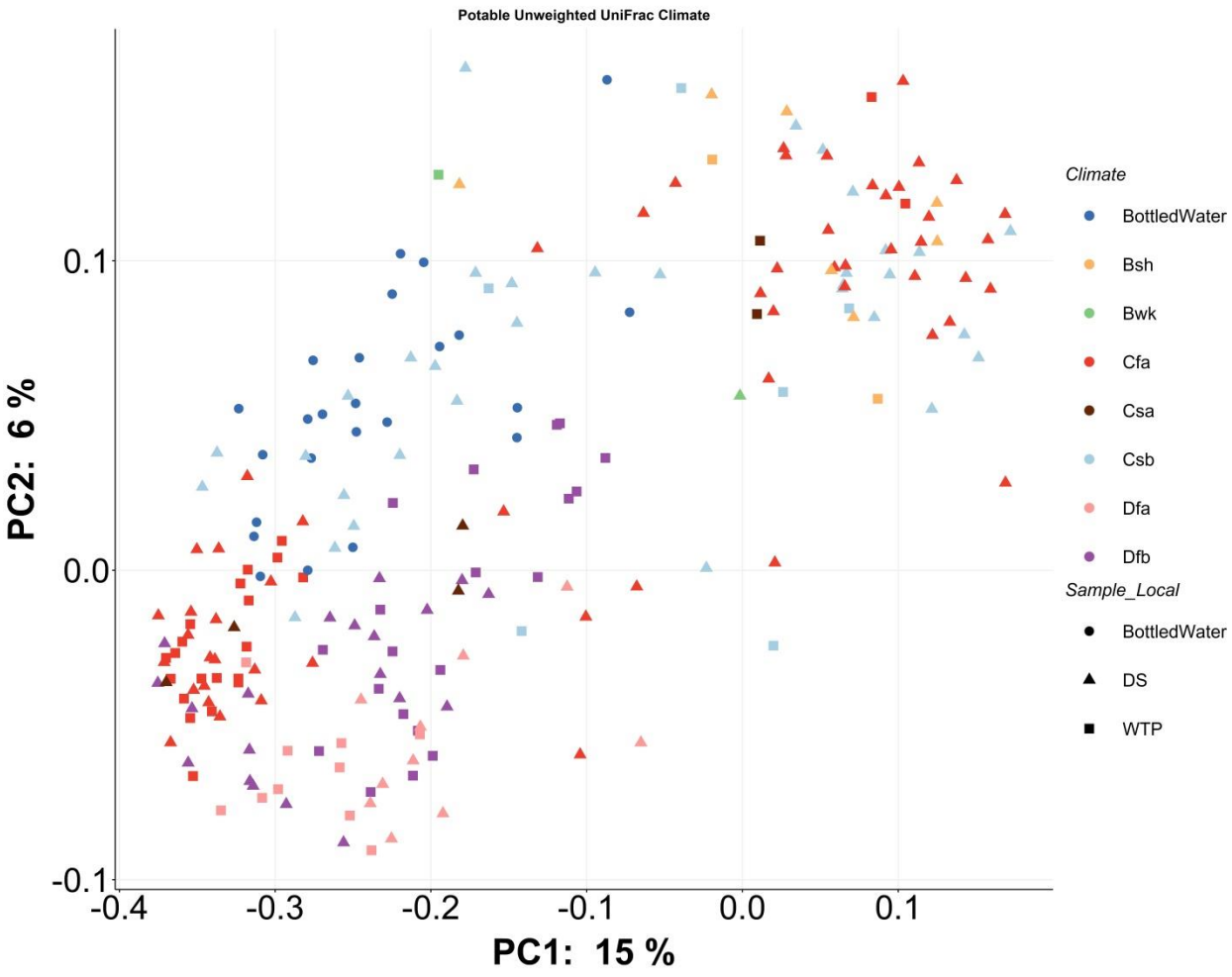
SI Figure G.5: Unweighted UniFrac beta diversity plot for all bulk water samples of interest, classified by intended water use and Köppen climate designator.



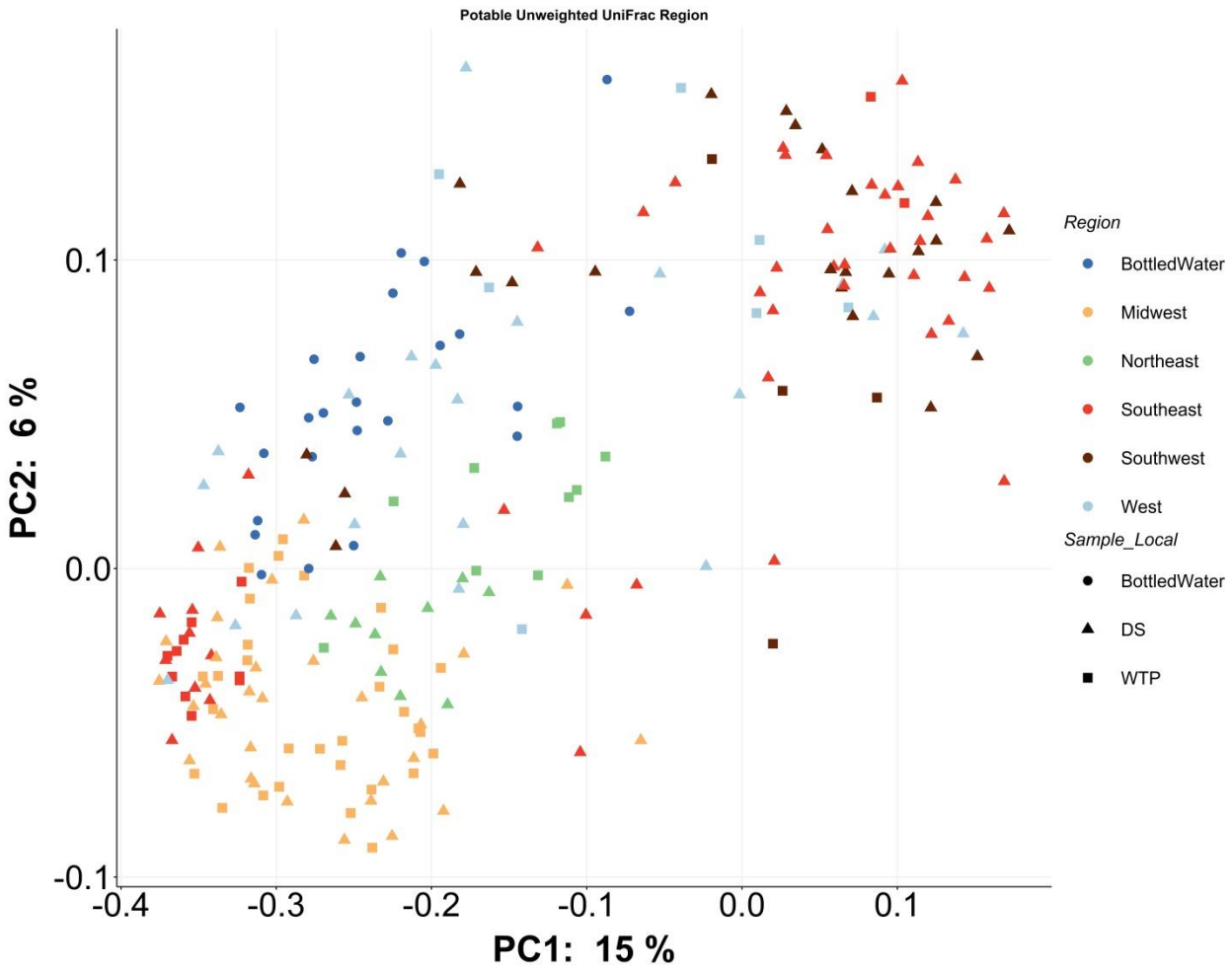
SI Figure G.6: Unweighted UniFrac beta diversity plot for all bulk water samples of interest, classified by intended water use and regional designator.



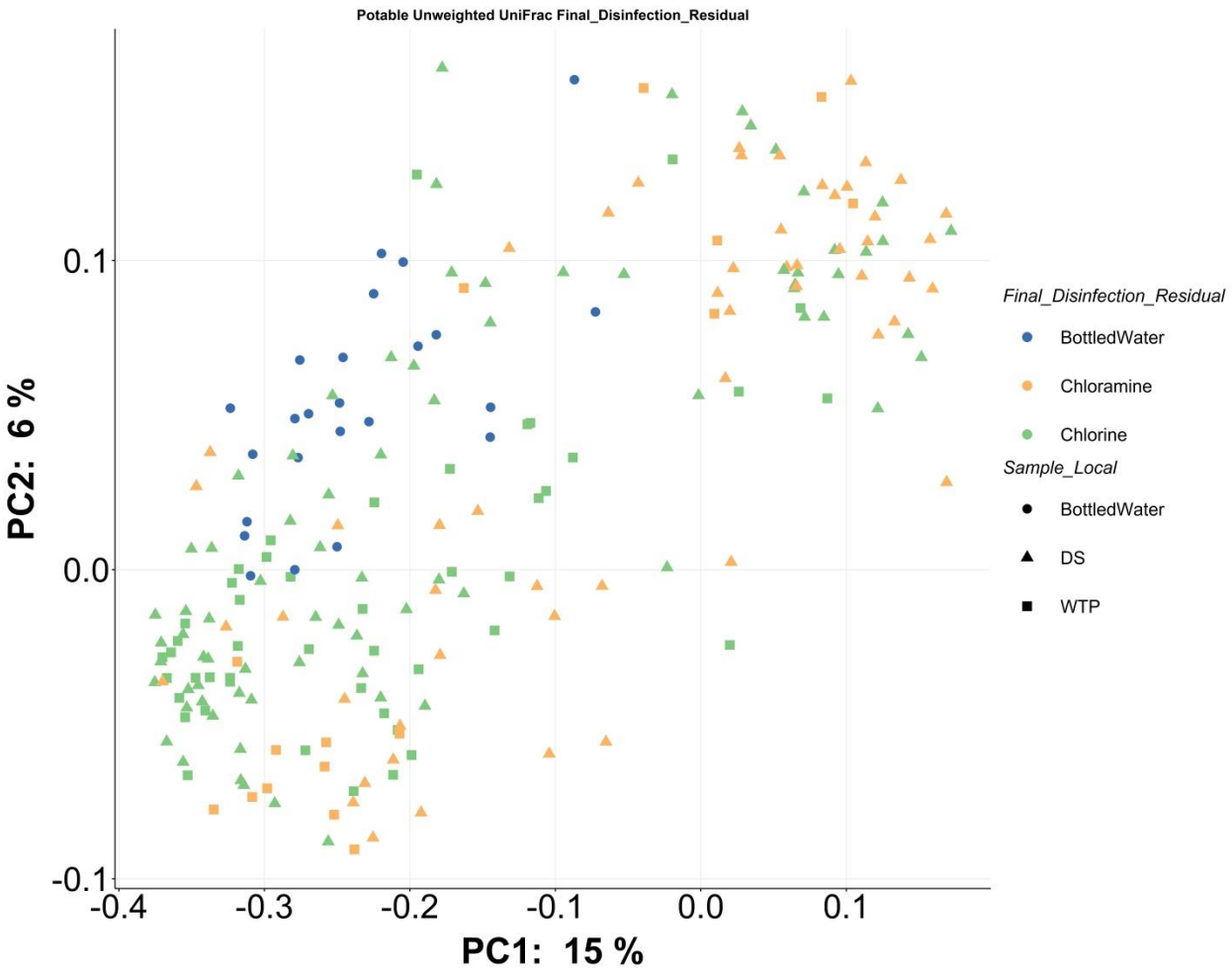
SI Figure G.7: Unweighted UniFrac beta diversity plot for all bulk water samples of interest, classified by DNA extraction kit.



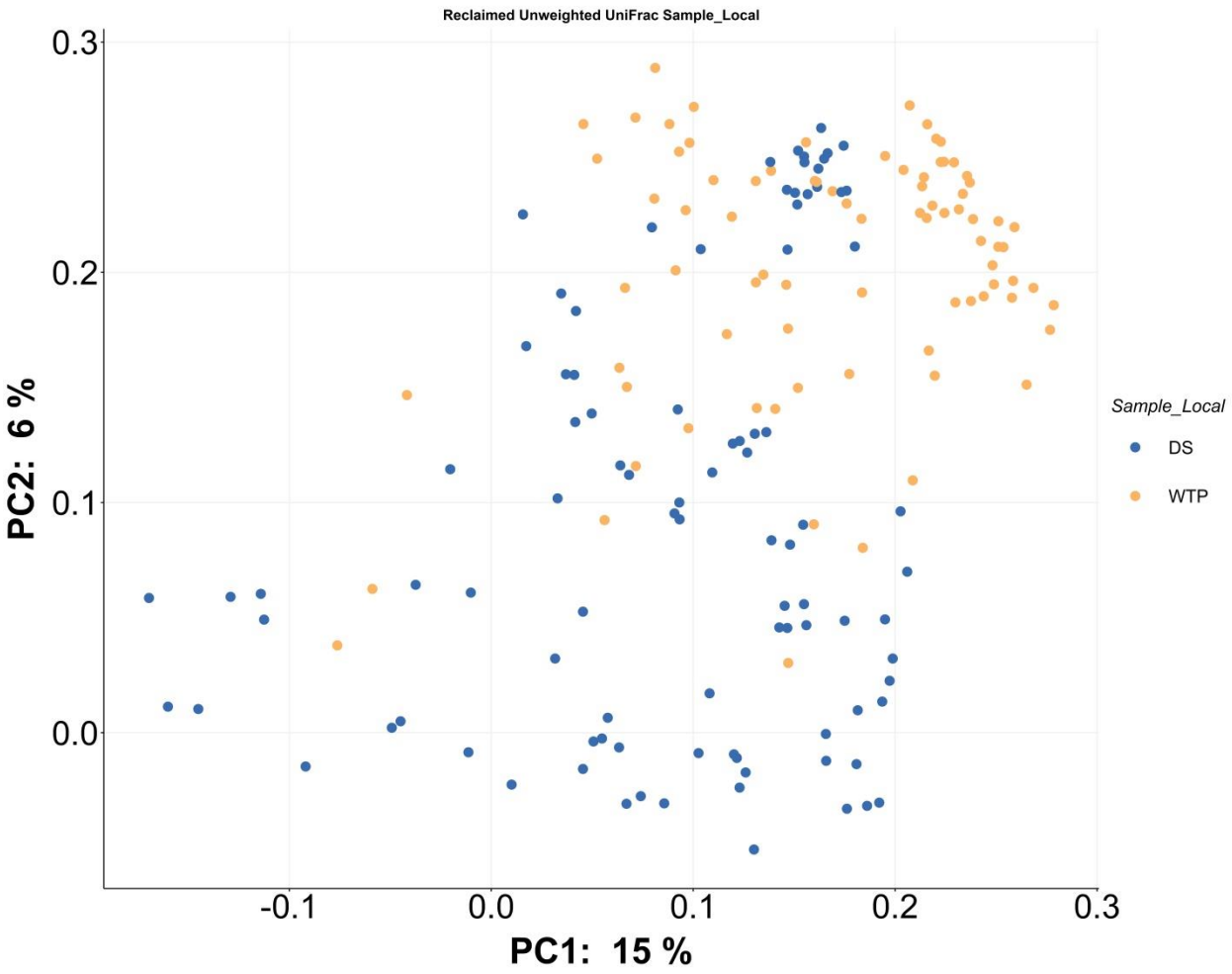
SI Figure G.8: Unweighted UniFrac beta diversity plot for all potable bulk water samples, classified by Köppen climate designator and sample location related to the POC or POU.



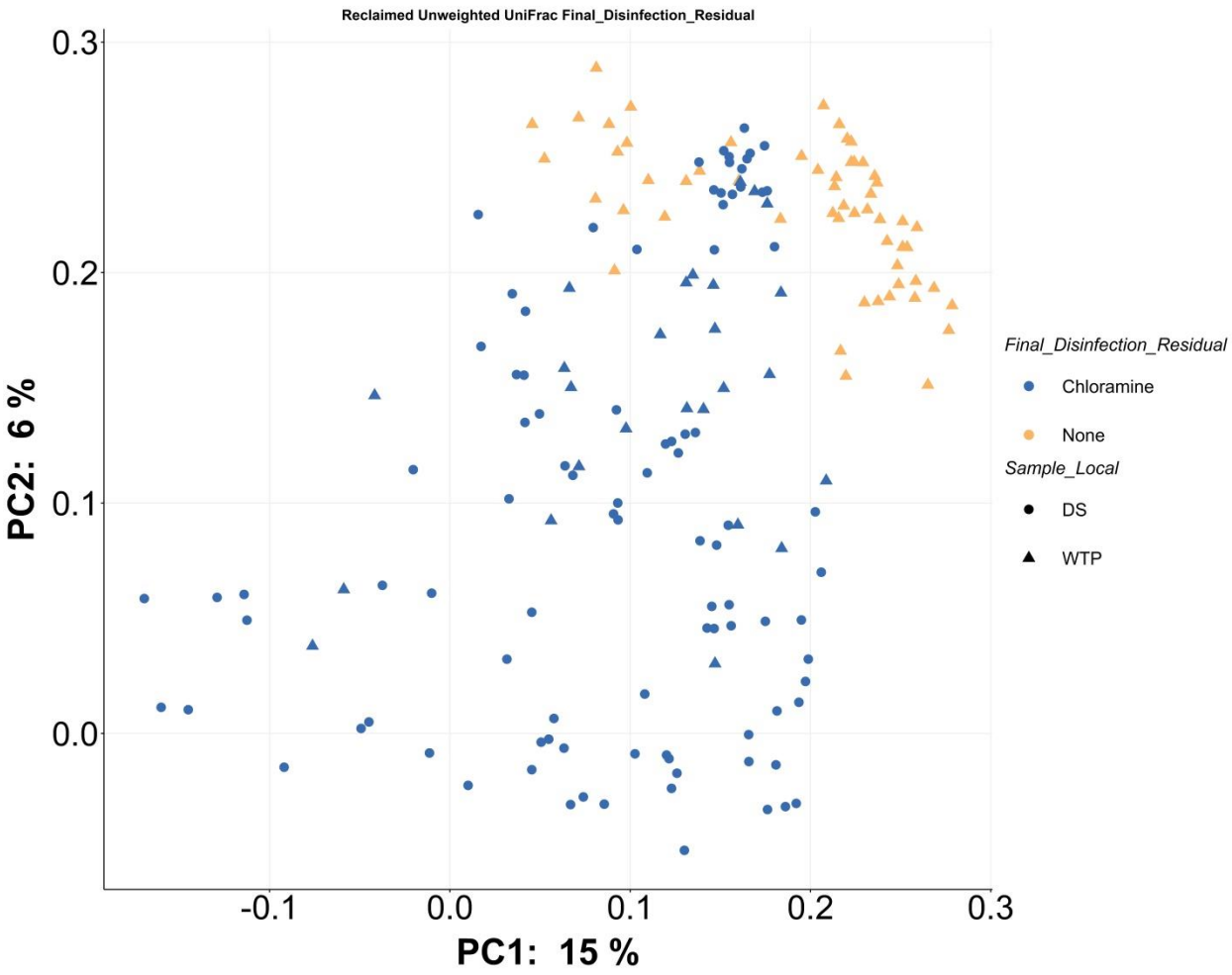
SI Figure G.9: Unweighted UniFrac beta diversity plot for all potable bulk water samples, classified by a regional designator and sample location related to the POC or POU.



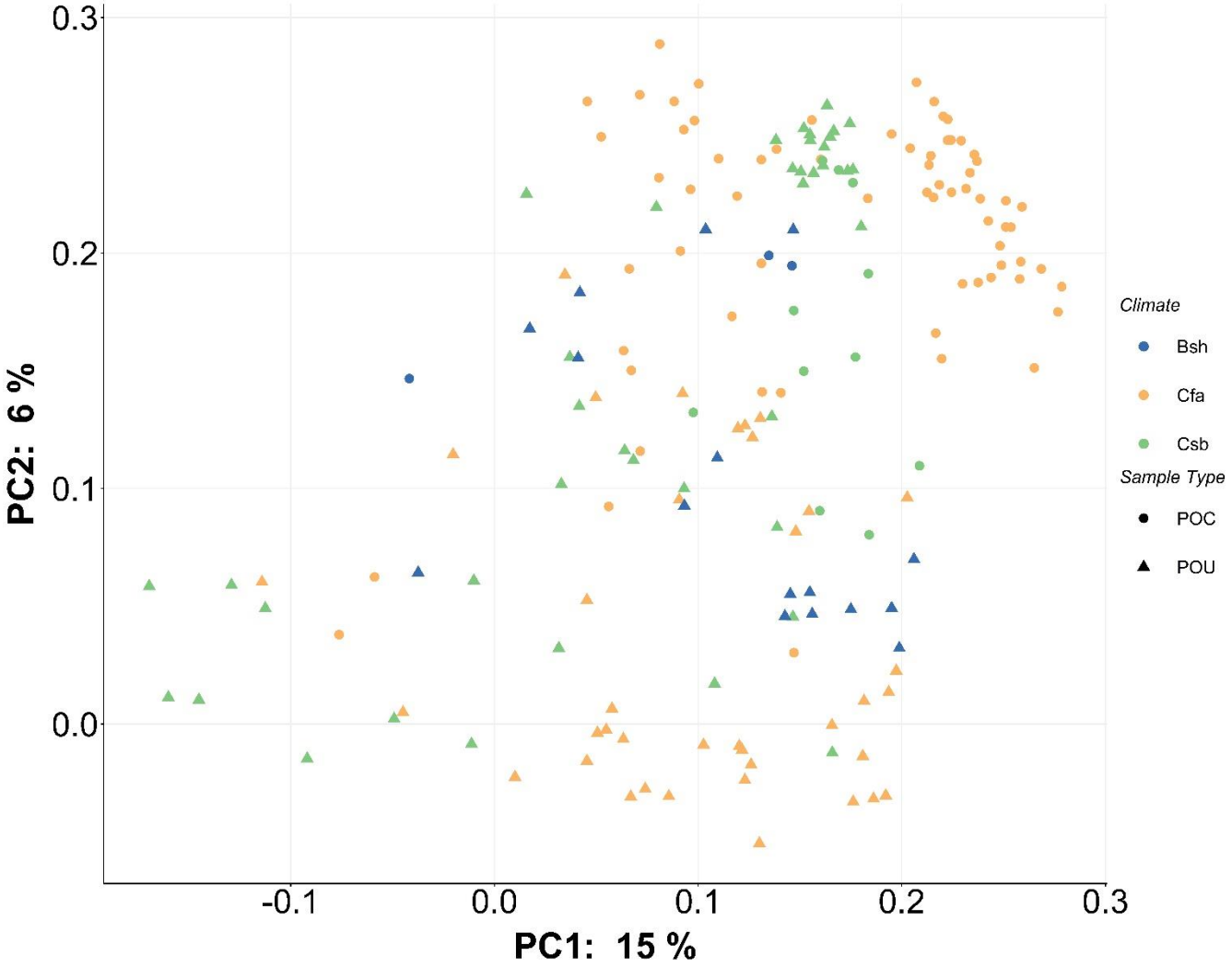
SI Figure G.10: Unweighted UniFrac beta diversity plot for all potable bulk water samples classified by final disinfection residual and sample location related to the POC or POU.



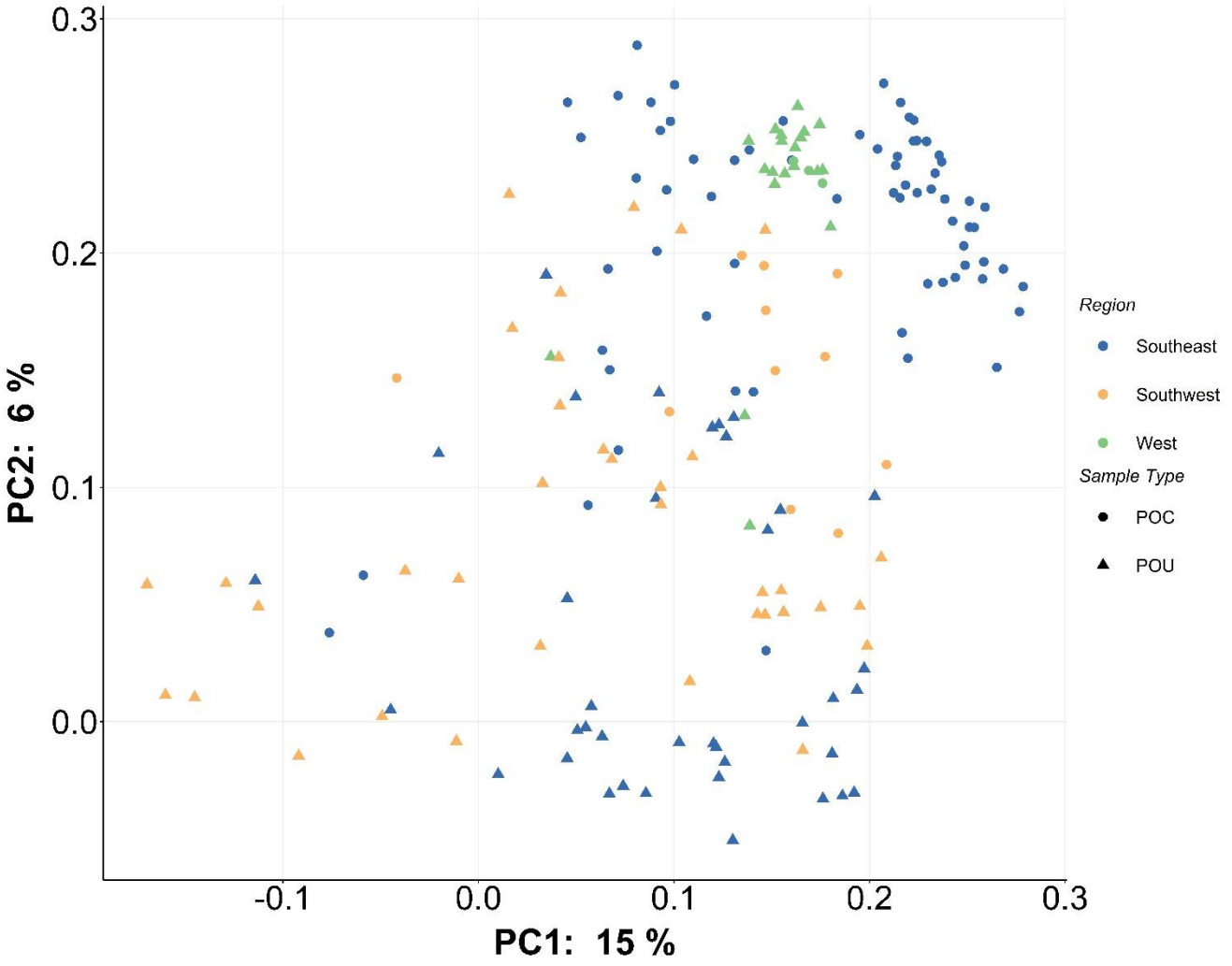
SI Figure G.11: Unweighted UniFrac beta diversity plot for all non-potable reuse water samples, classified by sample location related to the POC or POU.



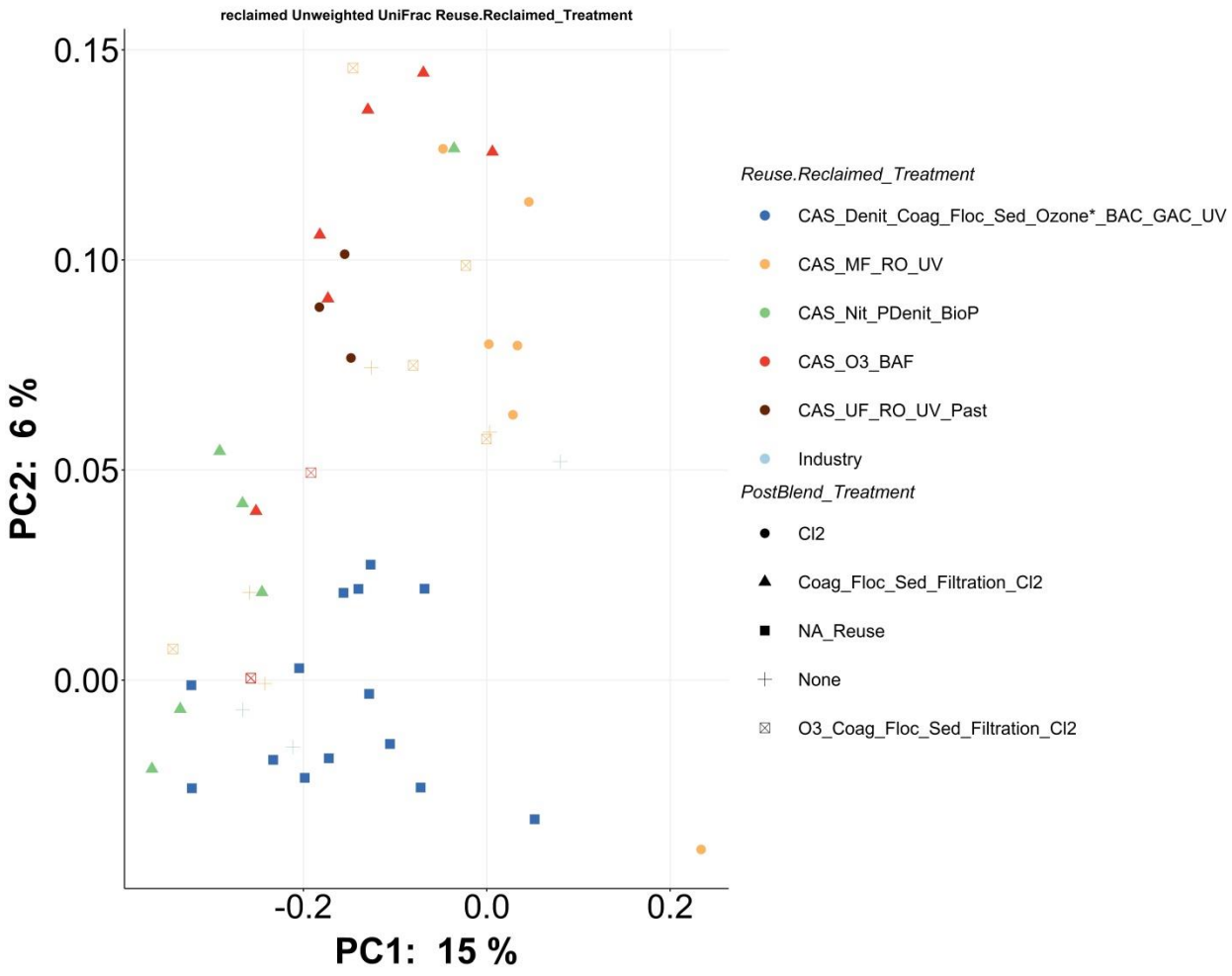
SI Figure G.12: Unweighted UniFrac beta diversity plot for all non-potable reuse bulk water samples classified by final disinfection residual and sample location related to the POC or POU.



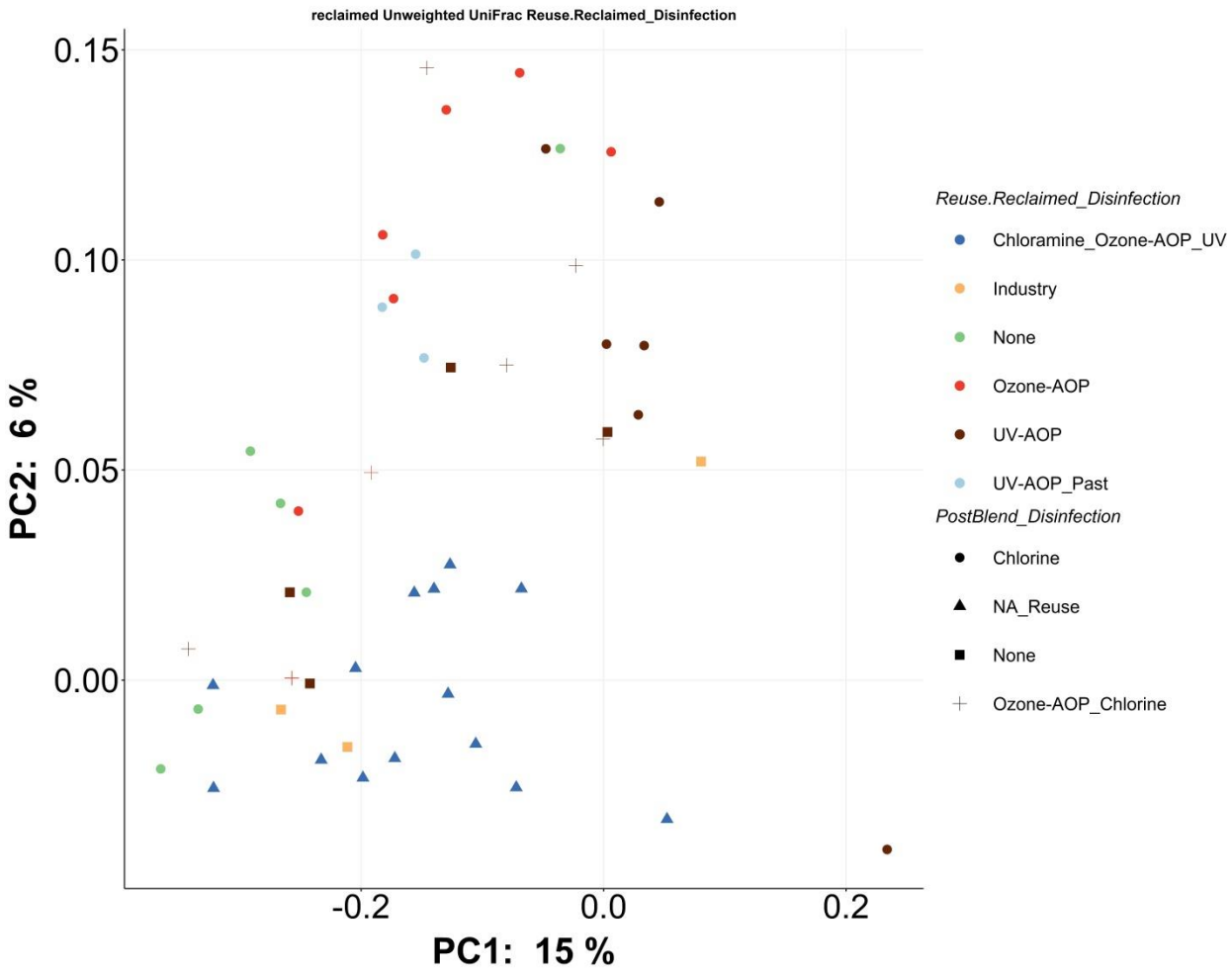
SI Figure G.13: Unweighted UniFrac beta diversity plot for all non-potable reuse water samples classified by Köppen climate designator and sample location related to the POC or POU.



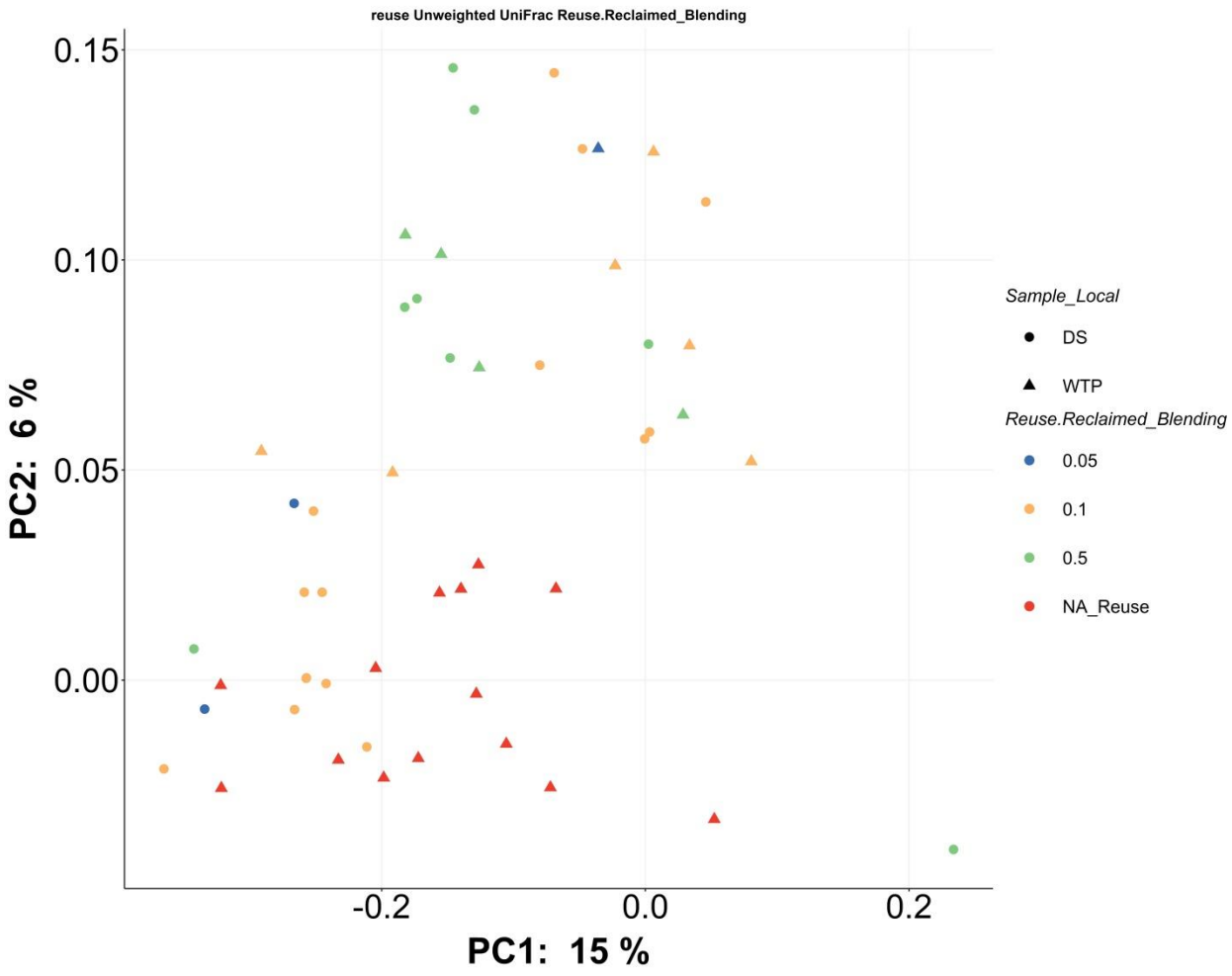
SI Figure G.14: Unweighted UniFrac beta diversity plot for all non-potable reuse water samples classified by region and sample location related to the POC or POU.



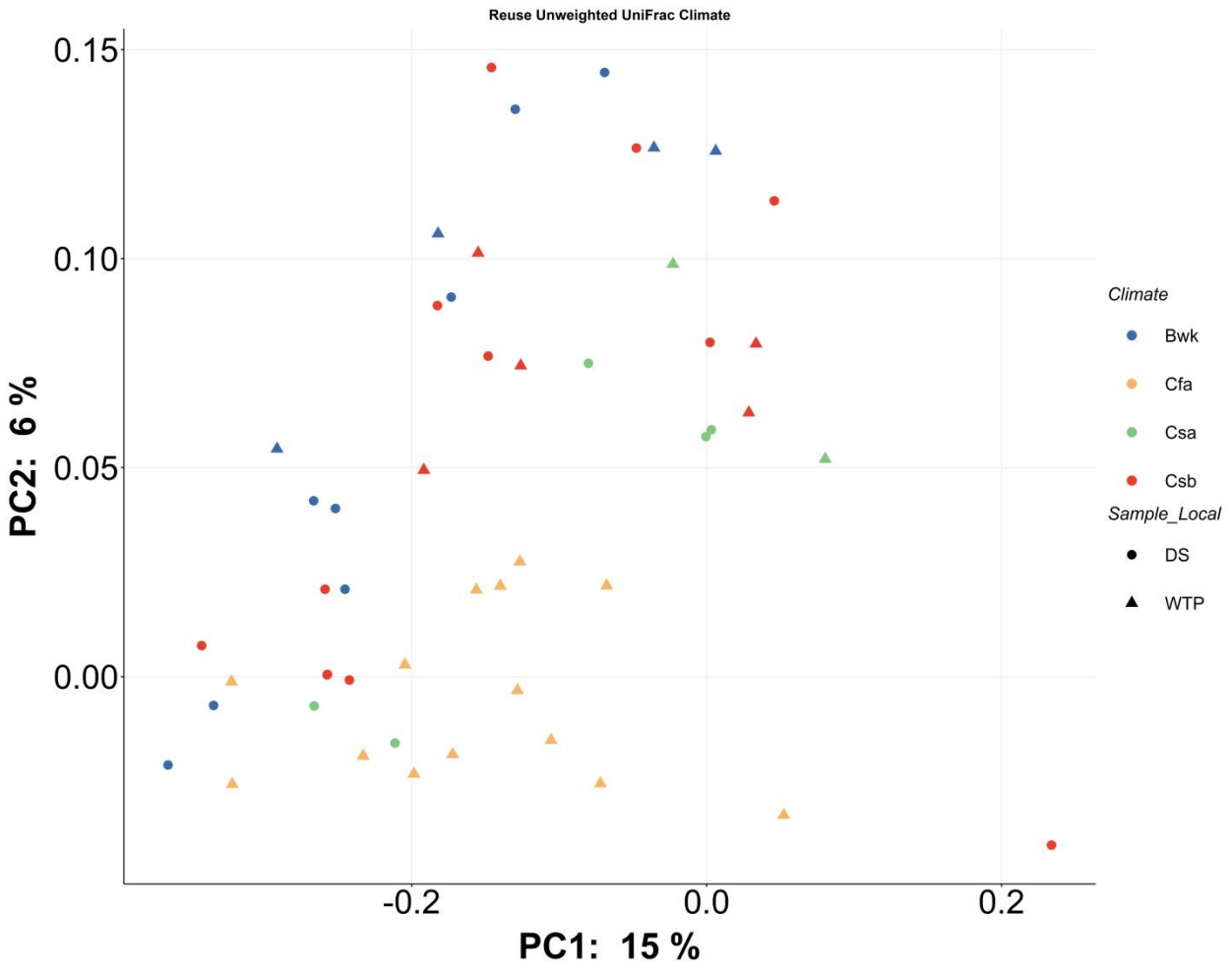
SI Figure G.15: Unweighted UniFrac beta diversity plot for all reuse bulk water samples, classified by treatment configuration before and after blending.



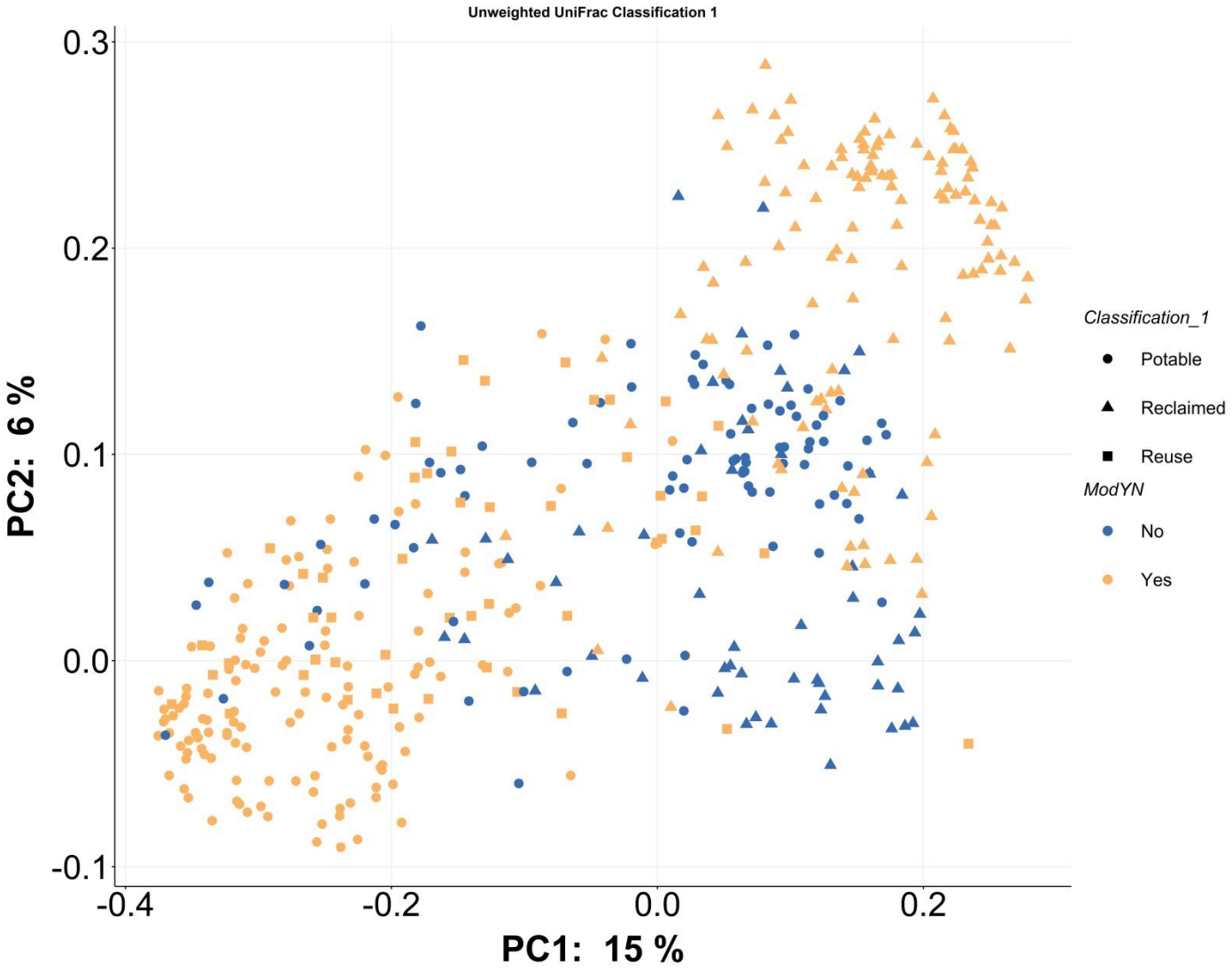
SI Figure G.16: Unweighted UniFrac beta diversity plot for all reuse bulk water samples, classified by disinfection processes before and after blending.



SI Figure G.17: Unweighted UniFrac beta diversity plot for all reuse bulk water samples, classified by blending ratio and sample location related to the POC or POU.



SI Figure G.18: Unweighted UniFrac beta diversity plot for all reuse bulk water samples, classified by Köppen climate designator and sample location related to the POC or POU.



SI Figure G.19: Unweighted UniFrac beta diversity plot for samples included in the modified core and discriminatory microbiome analysis, grouped by water type.

Subsection 3: OTU Based Core and Discriminatory Microbiome

In total, 970 unique OTUs were identified regardless of sample classification. When considering all samples, regardless of POC and POU distinctions, the potable water microbiome was found to have 3 core OTUs (identified in text by their first four characters; bcba, f480, and f898) all of which were derived from the phylum Proteobacteria with genera classification of *Sphingomonas*, *Ralstonia*, *Bradyrhizobium* respectively, SI Figure G.20 through SI Figure G.22 and SI Table G.14. These 3 core OTUs were similarly found when considering the POU subcategorization while only 2 (f480, and f898) were found to be core at the POC. Interestingly, OTU-bcba was only found in approximately 65% of potable POC samples while upwards of 96% at the POU, speaking to impact that distribution systems have on the selection of microbial communities. The core microbiome of potable reuse samples only constituted of OTUs that were also core to potable water, SI Figure G.20 through SI Figure G.22 and SI Table G.14. These

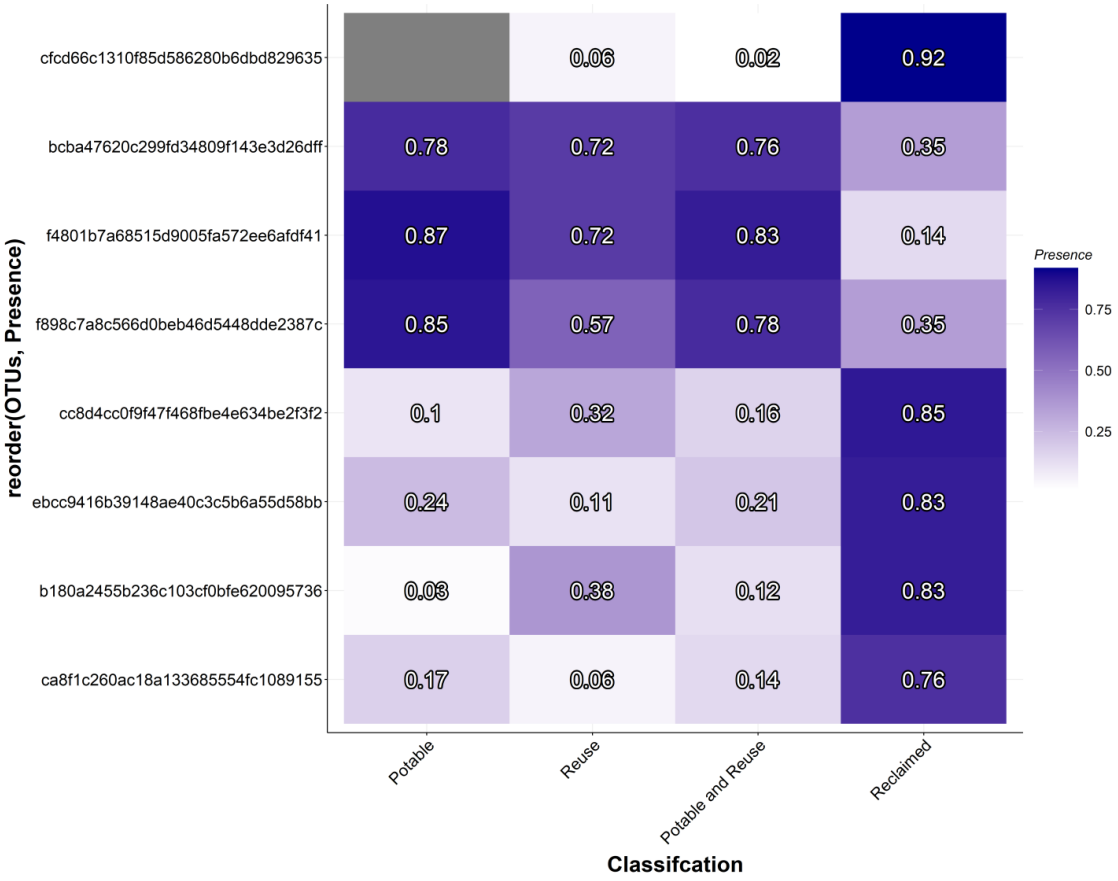
included the OTUs bcba and f480 which were found to be core when all samples were combined while only f480 was core at the POC and bcba was core at the POU. In all comparisons, there were no identifiable discriminatory OTUs between potable and potable reuse samples, supporting the idea of a core microbiome associated with the safe consumption of drinking water.

Samples intended for non-potable reuse were found to have more core OTUs than their potable and potable reuse counterparts with 5 OTUs (cfcd, cc8d, ebcc, b180, and ca8f) present within the entire dataset, 12 core OTUs at the POC (e6bc, e17c, d2a1, cfcd, ae7c, ebcc, b180, c97e, cc8d, e990, e79a, and b357), and 6 core OTUs at the POU (cfcd, a2f2, cc8d, adf8, ca8f, and a47f), SI Figure G.20 through SI Figure G.22 and SI Table G.14. This increase in core OTUs between potable and non-potable water uses can likely be attributed to either increased diversity among non-potable reuse samples, or less microbial variation following treatment intended for non-potable reuse. The latter is well documented by the decrease in core OTUs as produced waters are subjected to distribution systems and sampled at the POU where the non-potable reuse POU microbiome experienced more dispersion, and community variation as evident in the beta diversity analysis. When compared, the core OTUs among non-potable reuse POC and POU only share two OTUs that are independently considered core for each subcategorization. These included cc8d (taxonomically classified as the genus *Hydrogenophaga* from the phylum Proteobacteria) and cfcd (taxonomically classified as the genus *Flavobacterium* from the phylum Bacteroidetes) while ebcc, b180, and ca8f were found at high enough prevalence throughout both subcategory's to be consider core when both sample types were combined, but not independently.

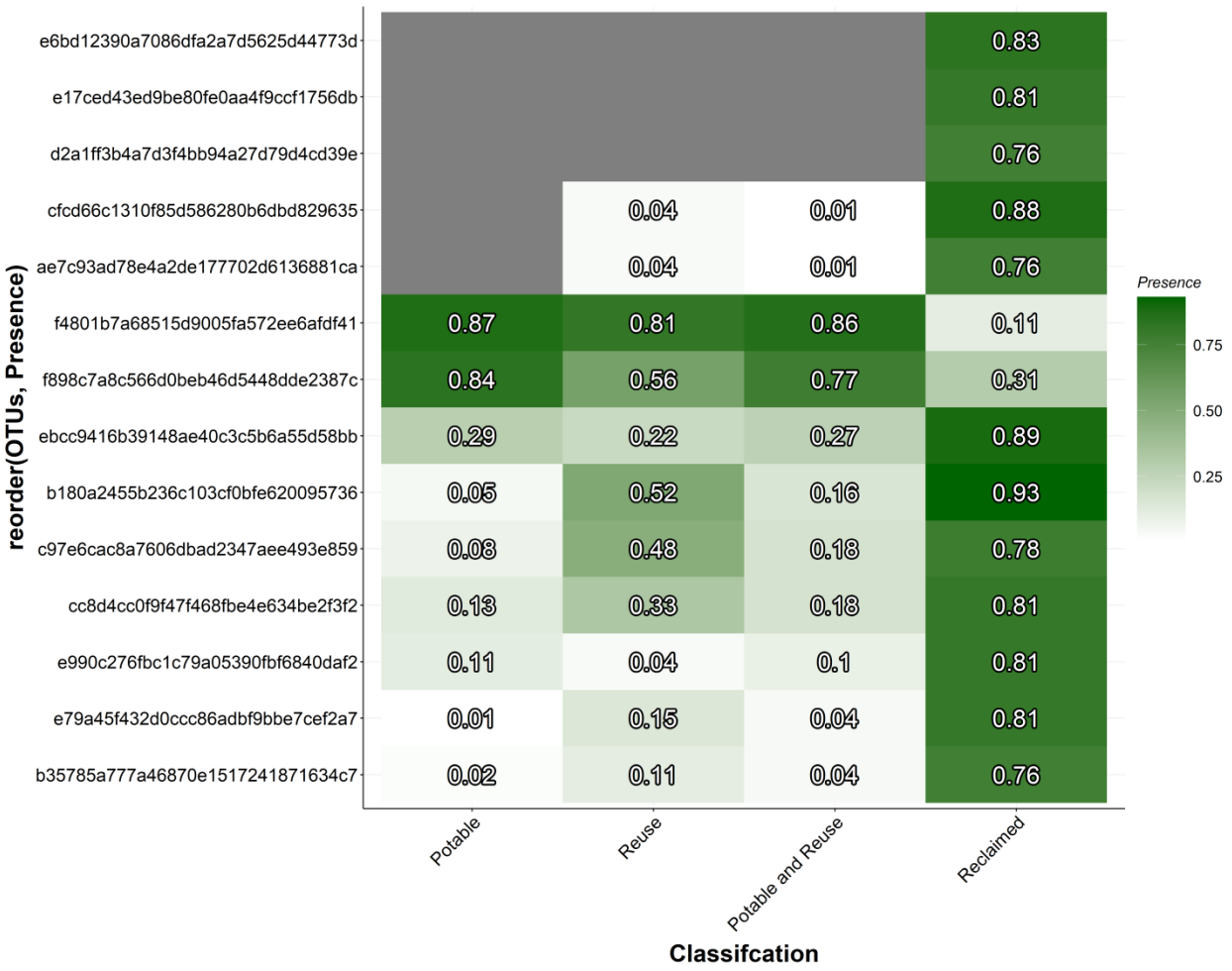
A comparison of core OTUs between the different water uses also allowed for the identification of discriminatory OTUs that were both core to one water use and uncommon to another. For the complete dataset, 4 core non-potable reuse OTUs were also discriminatory to potable and potable reuse samples (cfcd, cc8d, ebcc, and ca8f) with the OTU-cfcd being of particular interest due to its high presence in non-potable reuse waters (92%) and its very low presence in both potable (NA) and potable reuse waters (6%). Additionally, OTU b180 was found to be discriminatory between non-potable reuse and potable, but not potable reuse waters. From a core potable perspective, OTU f480 was found to be core to all potable and potable reuse samples and uncommon in all non-potable reuse systems.

Identified previously utilizing beta-diversity, the microbiomes of each intended water use was more divergent at the POC than at the POU, with the distribution system seemingly acting as a normalizing factor. This observation was supported by the core microbiome analysis with the identification of more discriminatory OTUs present when comparing samples from the POC (14 discriminatory OTUs) than when comparing samples from the POU (7 discriminatory OTUs), SI Figure G.20 through SI Figure G.22 and SI Table G.14. Of these 14 discriminatory OTUs directly after treatment, 9 were found to have been core to non-potable reuse samples and uncommon to both potable and potable reuse finished waters (e6bd, e17c, d2a1, cfcd, ae7c, ebcc, e990, e79a, and, b357) while 3 were core to non-potable reuse and uncommon to just potable waters (b180, c97e, and cc8d). From a potable perspective, OTUs f480 was core to both potable and potable reuse waters and uncommon in non-potable reuse systems while f89c was only core to potable systems and uncommon to non-potable reuse waters. At the POU, the number of discriminatory OTUs decreased to 7, with only the discriminatory relationship of cfcd shared

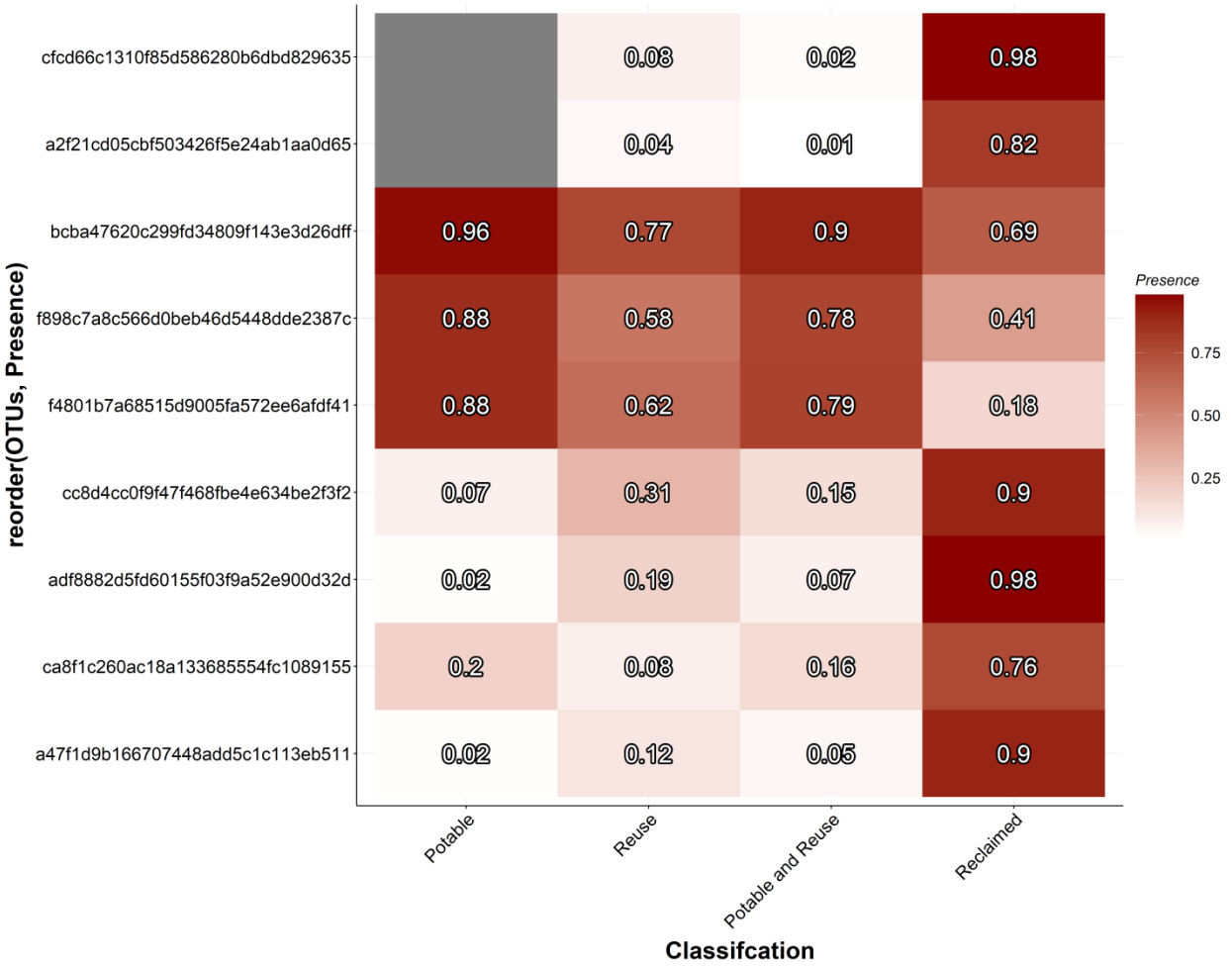
between non-potable reuse, potable, and potable reuse waters independently at both the POC and POU. Additionally at the POU, 5 OTUs were found to be core to non-potable reuse systems and uncommon in both potable and potable reuse waters (a2f2, cc8d, adf8, ca8f, and a47f) while 1 OTU was found to be core to potable distribution systems and uncommon in non-potable reuse distribution systems (f480).



SI Figure G.20: Unweighted core and discriminatory microbiome analysis conducted at the OTU level and grouped by intended water use. Average frequency of detection among all classified samples is presented. All bulk water samples, regardless of POC or POU distinctions.



SI Figure G.21: Unweighted core and discriminatory microbiome analysis conducted at the OTU level and grouped by intended water use. Average frequency of detection among all classified samples is presented. All bulk water samples at the POC.



SI Figure G.22: Unweighted core and discriminatory microbiome analysis conducted at the OTU level and grouped by intended water use. Average frequency of detection among all classified samples is presented. All bulk water samples at the POU. (C, Bottom Left) All bulk water samples at the POC.

SI Table G.14: Taxonomy reference key for identified core and discriminatory OTUs present in heat maps (A) through (C).

OTUs	Phylum	Class	Order	Family	Genus
b180a2455b236c103cf0bfe620095736	Firmicutes	Clostridia	Clostridiales	Peptostreptococcaceae	NA
bcba47620c299fd34809f143e3d26dff	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas
ca8f1c260ac18a133685554fc1089155	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rhodobacter
cc8d4cc0f9f47f468fbe4e534be2f3f2	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Hydrogenophaga
ebcc9416b39148ae40c3c5b6a55d58bb	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	NA
f4601b7a68515d9005fa572ee8afdf41	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Ralstonia
f898c7a8c566d0beb48c5448dde2387c	Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium
cfcd66c1310f85d586280b6cbd829635	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
b35785a777a46870e1517241671634c7	Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira
c97e6cac8a7606dbad2347aee493e559	Proteobacteria	Betaproteobacteria	Methylophilales	NA	NA
e79a45f432d0cac86adbf9bbe7cef2e7	Proteobacteria	Betaproteobacteria	Hydrogenophilales	Hydrogenophilaceae	Thiobacillus
e990c276fcc1c79a05390fbf6840daf2	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	NA
ae7c93ad78e4a2de177702d6136881ca	Chlorobi	SJA-28	NA	NA	NA
d2a1ff3b4a7d3f4bb94a27d79d4cd39e	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	Paludibacter
e17ced43ed9be80fe0aa4f9ccf1756db	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
e6cd12390a7086cfa2a7d5625d44773d	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	NA
a47f1d9b166707448add5c1c113eb511	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
adf8882d5fd80155f03f9a52e900d32d	Proteobacteria	Gammaproteobacteria	Alteromonadales	[Chromatiaceae]	Rheinheimera
a2f21cd05cbf503426f5e24ab1aa0d65	Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae	Azospira