Spring 2019

A Comparison of Aquatic Microbial Community Structure in a Rural Stream (Sawkill in Red Hook, NY) and an Urban Stream (Sawmill in Yonkers, NY)

Emlyn A. Ellerby
Bard College

A Comparison of Aquatic Microbial Community Structure in a Rural Stream (Sawkill in Red Hook, NY) and an Urban Stream (Sawmill in Yonkers, NY)

Senior Project Submitted to

The Division of Sciences, Mathematics, and Computing

of Bard College

by

Emlyn Ellerby

Annandale-on-Hudson, NY

A Comparison of Aquatic Microbial Community Structure in a Rural Stream (Sawkill in Red Hook, NY) and an Urban Stream (Sawmill in Yonkers, NY)

Senior Project Submitted to

The Division of Sciences, Mathematics, and Computing

of Bard College

by

Emlyn Ellerby

Annandale-on-Hudson, NY

A Comparison of Aquatic Microbial Community Structure in a Rural Stream (Sawkill in Red Hook, NY) and an Urban Stream (Sawmill in Yonkers, NY)

Senior Project Submitted to

The Division of Sciences, Mathematics, and Computing

of Bard College

by

Emlyn Ellerby

Annandale-on-Hudson, NY

Acknowledgements

I would like to thank everyone who helped make this project possible. My senior project board members for the feedback, guidance, and support they've given both in the classroom and during the progression of my project. Bard Water lab assistant Marco Spodek along with colleagues Azlan Maqbool and Renad Bdair for their assistance and companionship throughout. The Saw Kill Watershed Community for generously sharing their water quality data with me and enabling me to perform this study. My family and friends who have given great emotional support and encouragement throughout the process. The Environmental and Urban Studies department and Bard College for having given me this opportunity to conduct such research and open my eyes to the world of environmental science and ways in which we can make positive impacts on the environment. Lastly, and perhaps most importantly, thank you to my advisor and scientific mentor Elias Dueker for your tremendous helping effort over the course of completing this project. I am forever grateful for the direction, inspiration, and support you've given me during my time here.

Table of Contents

Abstract

Stream microbial communities serve as a foundation to higher trophic levels and aquatic organisms by utilizing nutrients and organic carbon in different ways, providing energy to be used in a constantly dynamic system. Variation in utilization of available resources enhances biodiversity and enables different bacteria to coexist. These aquatic microbial communities vary greatly across urban, suburban/rural, and pristine stream environments, as their composition can be significantly altered by both natural and anthropogenic activities. Anthropogenic land use and an increasingly urbanized landscape have been sources of pollution to waterways, releasing toxic contaminants that spread and can biomagnify through the food chain to higher trophic levels, disrupting essential biochemical and metabolic processes of aquatic organisms, particularly sensitive microbial communities. The microbial communities, in response, either develop resistance to the contaminant and survive at a limited functioning capacity, or succumb to the toxicity and slowly homogenise the microbial structure of a stream. An overloading of organic and inorganic contaminants can overwhelm the existing microbial communities and impair their ability to metabolise organic matter. Understanding how microbial communities change due to the presence of increasing pollutants is of fundamental importance to humans, particularly as these microbes are responsible for much waste absorption capacity in nature through decomposition of organic matter and reduction of toxic substances. A higher concentration of people in a city or a college campus can increase the loading of contaminants into nearby waterways through wastewater and sewage pipes. A high density of people situated in a rural setting can have a severe impact on the local aquatic environment due to wastewater effluent and pollutants collected in runoff. Comparing population gradients, observing their impacts on nearby stream microbial communities, and monitoring sources of pollution into waterways can serve to further our knowledge of how our waste influences these organisms so crucial to improving water quality. To address the question of differences in microbial community structure, I analyze the effects of external inputs and varying levels of urbanization on the aquatic microbial community structure of a Rural stream (Saw Kill in Red Hook, NY) and an Urban stream (Saw Mill in Harlem, NY), expecting a detectable change in microbial communities due to such external inputs.

1.0 Introduction

1.1 Stream Microbial Diversity

It is clear that global environmental change poses many potential threats to the structure and function of all aquatic ecosystems (Malmqvist and Rundle, 2002; Dudgeon et al., 2006; Zeglin, 2015), and that changing environmental factors at the watershed scale directly impact the biological function of lotic[1] ecosystems (Likens et al., 1970; Hynes, 1975; Mulholland et al., 2008; Palmer and Febria, 2012; Zeglin, 2015). Stream ecosystems are the primary receivers of nutrient and organic carbon exported from terrestrial ecosystems, and streams heavily influenced by agricultural and urban land uses are expected to have elevated nutrient and organic matter concentrations (Qu et al., 2017). Four-fifths of all the water we use comes from streams, and because of their rapid renewal rate as compared with lentic ecosystems, streams are the best single measure of available water supply (Christopherson 198, 1998). Rivers and streams contain approximately 0.006% of the freshwater available on Earth, and more than 30% of the renewable freshwater available for consumption is used for agricultural, industrial and domestic purposes, and the direct consequence of such usage (apart from increasing water scarcity) is large amounts of waste being injected into waterways, modifying the structure and composition of rivers and streams by altering their geomorphology, temperature, pH, nutrient availability, and aquatic biotic community (Medeiros et al., 2016).

---

[1] Lotic — Rapidly moving freshwater riverine environments characterized by unidirectional flow, continuous physical change, and a high degree of spatial and temporal heterogeneity (Caspers, 1979). Lotic environments are characterized by relatively high levels of dissolved oxygen due to the turbulence of flowing waters and include both intermittent and perennial streams (McCafferty, 1998; Astrachan, 2017).

Of fundamental importance to the aquatic biotic community is the ability of microbes to decompose organic matter, often in the form of plant litter that can be digested by various microbial communities, utilizing the released nutrients from organic matter decomposition to reproduce (and in the case of anaerobic bacteria, remineralize organic nitrogenous substances back into forms of inorganic nitrogen such as nitrate (NO3-) or nitrite (NO2-)), and guarantee future ability of microbial offspring to carry out these specialized, decompositional processes. Both terrestrial and aquatic microbial communities have evolved for billions of years prior to the arrival of *homo sapiens* on Earth, and such extensive evolution and interactions with a constantly dynamic environment has allowed for the emergence of an array of diverse bacteria and fungi, capable of carrying out highly specialized orders of decomposition and the ability to breakdown substances that certain other bacteria cannot (Rosenberg et al., 2016). A trade-off exists, and competition increases for available resources and nutrients, particularly Carbon (C), Nitrogen (N), Phosphorous (P), Sulfur (S), enabling certain members of the microbial community with competitive advantage to flourish and outcompete others, helping shape the local and regional microbial gene pool (Hoostal et al., 2008).

1.2 Climate Change

As the effects of global warming accelerate climate change, releasing vast amounts of water into the sea and atmosphere, charged water molecules are released that can exacerbate the intensity of hurricanes and other natural disasters, more pressure builds in the atmosphere and in the oceans, and the hydrological cycle that is influential to many of Earth's biogeochemical processes becomes stressed and unpredictable (Allan, 2004; Feng et al., 2015; Findlay 2010; Wilhelm at al., 2013; USGS; NOAA; Union of Concerned Scientists (UCSUSA)). Increasing

the magnitude of free radicalized water molecules will cause more precipitation events and in turn, more runoff, which can dictate the stream microbial communities along the run of the tributary (Ancion et al., 2012). Storm events are unique in that exotic bacteria can potentially come into contact with a previously novel environment by using turbid floodwaters, full of suspended solid particles, as a medium of transport, thus enabling dispersal and possible emergence of novel bacteria capable of resisting various contaminants (Benoit et al., 1999; Harding, 2005; Zeglin, 2015).

Additionally, increasing precipitation events will generate more runoff over impervious surfaces providing foundation to urban regions, of which a multitude of inorganic contaminants are washed into nearby waterways (Hosen et al., 2017). Often these contaminants are recalcitrant and cannot be removed through wastewater treatment processes (WWTP) involving secondary and tertiary treatment, and these inorganic pollutants can be very toxic to a wide range of organisms from the microbial scale to humans and larger animals exposed through contact or drinking (Münze et al., 2017; Rosi et al., 2018; Zeglin, 2015).

As increasing anthropogenic activity and industrialization accelerate the effects of climate change, no where is the evidence more clear as the cryosphere, where ice caps and glaciers are melting at an alarming rate (Wilhelm et al., 2013; National Snow and Ice Data Center (NSIDC)). Ice ecosystems can harbour complex microbial communities, with some of the most abundant bacterial phyla detected in glacier ice being *Proteobacteria, Actinobacteria, Bacteroidetes, Firmicutes,* as well as photosynthetic *Cyanobacteria. Proteobacteria, Actinobacteria and Bacteroidetes* tend to be more dominant phyla in ice, streamwater and biofilms, whereas *Cyanobacteria*/chloroplasts are more abundant in glacial ice and biofilm

communities (Wilhelm et al., 2013). This same study by Wilhelm et al. found that there was low commonality of taxa both at the OTU level and family level between ice and biofilm communities, suggesting minor contribution of the ice communities to the biofilm assembly in the glacier-fed streams and further emphasizing the importance of preserving the diversity of glacial microbial communities as they are unique and potentially nonexistent elsewhere (Wilhelm et al., 2013). As glaciers recede, they change the hierarchical habitat template of the riverine landscape where environmental processes operating at local and regional scales differentially affect life in the glacier-fed streams (Sigler et al., 2002;Wilhelm et al., 2013). Melting glaciers mobilise ice-locked organic matter with implications for downstream carbon cycling and heterotrophic activity (Sigler et al., 2002;Wilhelm et al., 2013), and as the cryosphere continues to melt from global warming, glacial retreat may contribute to the homogenisation of microbial communities among glacier-fed streams (Wilhelm et al., 2013).

1.3 Water Quality

Anthropogenic land-use changes can increase nutrient inputs into streams, enhancing autotrophic production (Boëchat et al., 2009) and reducing the aquatic-riparian connectivity through reduction of input of terrestrial coarse particulate organic carbon that fuels aquatic heterotrophic production (Gregory et al., 1991; Campbell et al., 1992), and in turn, changes in the nutritional quality of aquatic communities can be expected to affect ecosystem processes, such as energy flux, secondary productivity, and ecosystem metabolism (Müller-Navarra, 2008). Microbial biofilms, which are mixed assemblages of microorganisms such as bacteria and algae encased in an extracellular matrix and attached to a surface, are ubiquitous in aquatic habitats and are frequently both structurally and taxonomically complex (Rosi et al., 2018). They are key

components of stream ecosystems because they are major drivers of nitrogen and carbon cycling (Battin et al. 2003), they are the base of stream food webs (Hall and Meyer 1998, Pusch et al. 1998), and they contribute to important ecosystem services, such as decreasing nutrient pollution and bioremediation of organic pollutants (Rosi et al., 2018). A number of external factors, often consequences of anthropogenic activities, can influence the assemblage of these microbial communities and biofilms greatly (Lyautey et al., 2005; Qu et al., 2017).

1.4 Urban and Rural/Agriculture

The urbanization phenomenon parallels increasing human population density, and as previously uninhabited lands become settlements, a supply of fresh water is needed. Urbanization not only alters the instream habitat, chemistry and flow regime, but also fragments terrestrial habitat necessary for the movement and reproduction of stream autotrophs and invertebrates, affecting the overall net primary productivity (Urban et al., 2006). Industry, mining, and urbanization have considerably increased metal concentrations in surface waters, with urban dominated catchments generally having trace metal concentrations several times higher than background levels which may result in significant damage to ecosystems (Ancion et al., 2012). Benthic macroinvertebrates have been widely used to evaluate the ecological impacts of metal contamination in streams, as metal contamination can reduce benthic macroinvertebrate species richness, density, growth, and production (Maret et al., 2003; Gray and Delaney, 2008), and while the effects were compounded with factors such as altitude, temperature, stream width, turbidity, and heavy metals, the influence of heavy metals on the benthic communities was clear, a correlation showing total abundance and species richness decreasing with higher heavy metal concentrations (Qu et al., 2010). Additionally, benthic macroinvertebrates were found to be

more sensitive to heavy metals at higher elevations (Clements et al., 2000; Qu et al., 2010), possibly due to lower levels of oxygen which aerobic bacteria utilize in respiration/decomposition processes. Streams less exposed to increasing urbanization are a more sensitive aquatic environment than those streams previously exposed, where microbial communities show higher levels of resistance (Qu et al., 2010). This was the case in a study conducted by Rosi et al. when observing effects of varying concentrations of pharmaceuticals and drugs in stream environments, with the results indicating that urban streams had higher concentrations of drugs and contained biofilms with greater functional resistance to drugs than biofilms from less urban streams (Rosi et al., 2018). Additionally, however, biofilms in urban streams had significantly altered community composition in response to drugs, showing that certain contaminants can influence and shape the structure of microbial communities even in already polluted environments (Rosi et al., 2018).

With higher concentrations of heavy metals in urban areas than more suburban/rural areas (Fang et al., 2011; Szynkowska et al., 2009), microbial communities from polluted regions can demonstrate greater metal tolerance and be more resilient to inoculations with moderate concentrations of heavy metals than microbial communities from non-polluted regions, as shown by enzymatic activities of hydrolases from microbial communities in both polluted and unpolluted sediments of Lake Erie, with an increase in enzymatic activity in the polluted sediments indicating these communities are more resilient to heavy metal stress (Hoostal et al., 2008). Bacteria have adapted multiple heavy metal tolerance mechanisms, which can be capable of being spread through a bacterial community by lateral gene transfer, therefore heavy metals may act as important selective agents driving the evolution of microbial communities (Hoostal et

al., 2008). Urbanization landscapes are likely the sources of novel microbial taxa not found in undisturbed stream ecosystems, including taxa from sewage and septic systems, water distribution systems, and stormwater management ponds (Hosen et al., 2017).

Microbial communities present in marine sediments primarily decompose organic matter derived from plant litter but also play a pivotal role in the transformation of pollutants (Yao et al., 2016). Where they occur, heavy metals are pollutants of considerable concern because they are not usually eliminated from aquatic ecosystems by natural processes, but rather accumulated in sediments or biota, or transported to other ecosystems (Harding, 2005). Concentrations of metals and organometallic compounds in natural habitats may be reduced by microbial action, with these transformations capable of being carried out in a wide variety of habitats including lake and river sediments, soil, river water, and activated sludge, and in each place the microbial composition has been significantly different (Gadd and Griffiths, 1978; Hoostal et al., 2008; Igiri et al., 2018; Vishnivetskaya et al., 2011). For instance, several types of bacteria and yeast have been shown capable of the oxidation of elemental mercury to its cationic form, including *E. coli, Pseudomonas fluorescens, P. aeruginosa, Citrobacter* sp., *Bacillus subtilis,* and *B. megatherium* (Gadd and Griffiths, 1978).

Metal ions bind easily with suspended particles such as silt or organic matter (Ancion *et al.*, 2012), and significantly, the amount of dissolved organic matter (particularly in eutrophic or anthropogenically polluted systems) can also influence the toxicity of various metals, as exemplified by streams on the West Coast of the South Island of New Zealand (Harding, 2005). High concentrations of dissolved organic carbon (predominantly humic and fulvic acids derived from decomposing vegetation, giving these streams a brown color) greatly reduces the toxicity of

aluminum to benthic macroinvertebrates and fish, as it adsorbs to and is complexed with dissolved organic matter (Collier *et al.,* 1990; Harding, 2005).

Trace amounts of certain metals, including sodium, potassium, magnesium, calcium, manganese, iron, cobalt, copper, zinc, and molybdenum, are present in varying concentrations in all living tissues and are essential if an organism is to grow and metabolise successfully (Harding, 2005; Kelly, 1988), and the subsequent loss or removal of trace metals from an organism results in impaired biological functioning, while overexposure will have toxic effects (Harding, 2005; Hoostal et al., 2008; Liess et al., 2017; Qu et al., 2010; Rosenberg et al., 2016; Rosi et al., 2018). More toxic metals such as arsenic, cadmium, chromium, copper and mercury frequently accumulate in aquatic plants and in river and lake sediments, and some of these elements can be re-mobilised and incorporated into food webs, and can affect the physiology, growth and reproduction of organisms at multiple trophic levels (Kelly, 1988; Harding, 2005) through bioaccumulation, thus emphasizing the importance of microbial communities in their role as bioremediators and toxicity reductors.

Agricultural land use degrades streams by increasing nonpoint inputs of pollutants as well as higher input of sediments, nutrients, and pesticides, impacting riparian and stream channel habitat and altering flow (Allan, 2004). Heavy metal contamination from industrial inputs and increasing urbanization provide for more point source inputs of inorganic pollutants and can encourage heavy metal tolerance among the microbial communities (Urban et al., 2006). However, if certain members of the microbial community cannot develop resistance to the same extent as the rest of the community, their contribution to the community will be lost, and the

ability of the community to perform the above-mentioned processes will be impaired (Ross et al., 1989).

While heavy metal contamination is a common source of inorganic pollution to streams and lotic waterways, organic pollutants in the form of nitrogenous and phosphoric compounds have created a worldwide problem by providing excess nutrients to lentic ecosystems, and causing eutrophic conditions and the formation of harmful algal blooms (HABs)(Qu et al., 2017). Concentrated animal feeding operations (CAFOs) seep massive amounts of animal feces and waste into nearby streams and eventually lentic waters, where eutrophic conditions can ensue due to nutrient overloading (Hosen et al., 2017; Utz et al., 2016). Agriculture and urban use increase NO3−, NH4 +, and SRP concentrations as a result of excess petroleum-based nitrogenous fertilizer application, as well as sewage and septic inputs, which is carried by runoff into nearby bodies of water, reducing water quality and altering stream communities (Medeiros et al., 2016; Qu et al., 2017).

A fundamental difference with such organic pollution is that while it may not be immediately toxic to the environment through which it passes, it increases the readily available nutrients, and this may selectively stimulate some species or microbial groups, whose increased growth displaces potential competitors (Sridhar et al., 2001), leading to the possible homogenisation of the microbial community structure and lesser ability for microbial metabolism and OM decomposition.

Seasonal fluxes in nutrient enrichment of waterways are commonly evident, as exemplified by a study conducted by Feng et al. of Chesapeake Bay nitrogen fluxes derived from a land-estuarine ocean biogeochemical modeling system. A diverse assemblage of

phytoplankton species is responsible for the high rates of primary production observed in the bay, although in general, the timing, position, and magnitude of the spring bloom is determined by the high fluxes of riverine dissolved organic nutrients entering the bay, whereas the mean and variability of summertime phytoplankton concentrations are determined more by the degree of nutrient regeneration. Hypoxic (and eventually anoxic) conditions can occur in estuarine subpycnocline waters with a lack of aeration (Feng et al., 2015). During such periods, remineralization of organic matter in the water column transitions from an aerobic to an anaerobic process via facultative anaerobes that shift to alternative electron acceptors such as nitrate or nitrite (Feng et al., 2015). Such water column denitrification has previously been considered in marine ecosystem models of other hypoxic systems, such as the Black Sea and the Arabian Sea, and classifies the Chesapeake Bay as a net autotrophic estuary, as production of organic nitrogen exceeds the loss of organic nitrogen due to remineralization processes (Oguz et al, 2002; Resplandy et al., 2012; Feng et al., 2015).

An objective of this study is to analyze metagenomic[2] comparisons of relatively preserved and polluted areas of freshwater streams and how this may contribute significantly to a better understanding of the anthropogenic impacts on aquatic environments, as metagenomic analysis will provide information about the diversity and distribution of the different members of a microbial community and their metabolic potential. Stream microbial diversity is extremely important in that bacteria account for a significant portion of the transformation and use of organic matter in riverine systems, particularly their role as a nitrogen metabolism in streams and

[2] Metagenomic - The study of all genetic material from all organisms in a defined sample (Handelsman et al., 1998) Term used to describe a selection of tools and techniques that enable us to uncover DNA from the organisms in an environment (which can comprise any ecosystem, from soil to human intestinal tract) (Gilbert, 2013)

the denitrification of streams to protect downstream lakes and other lentic[3] environments from eutrophication[4] (Qu et al., 2017).

A study conducted by Wang et al. (2014) found that a diverse non-urban microbial community maintained denitrification potential in the face of multiple urban stressors (heavy metals, temperature, and elevated salt concentrations), whereas a less diverse urban microbial community did not, supporting the notion of greater stream microbial diversity in the non-urban stream.

1.4 Approaches for Saw Kill and Saw Mill

The Saw Kill watershed has a drainage area of 68 km^2 and includes a broad range of land use and land cover types (e.g., forest, wetland, cropland, transportation, residential, and commercial) along the length of the 14.3 mile stream (Zelewski et al., 2001). Most of the developed land in the watershed is associated with residential land use, with medium and high-density residential land use primarily located in the Village of Red Hook and nearby subdivisions, as well as Bard College in the lower portion of the watershed (Street et al., 2018). Commercial areas in the watershed are limited and concentrated primarily in the Village of Red Hook. Forests and wetlands comprise more than 60% of the watershed, with the remaining 24% of the watershed devoted to hay and pasture-related agricultural uses (Street et al., 2018). Within the area of the 100-year floodplain, 13.1% is developed, 18.5% is devoted to hay and pasture, while 64.5% is covered by forest, shrubland, or wetland (Street et al., 2018). The Saw Kill Watershed Community (SKWC) is a tightly knit group of community residents living in the local

[3] Lentic - Stationary freshwater such as lakes and ponds (Zeglin, 2015)

[4] Eutrophication - a body of water receives an excessive nutrient load, particularly phosphorus and nitrogen. This often results in an overgrowth of algae. As the algae die and decompose, oxygen is depleted from the water, and this lack of oxygen in the water causes the death of aquatic animals, like fish. (USGS)

area that care deeply for protecting the stream water quality; for clean drinking water, flood protection, water supply, recreation, and ecological resources. Bard College uses the Saw Kill as its source of drinking water, uptaking water just above the waterfall close to Bard's Water Treatment Plant, and also has a Wastewater Treatment Plant (WTTP) situated near the lower section of the stream towards the mouth, releasing treated effluent ~200 meters upstream from the mouth (Spodek, 2017) that. Additionally, a number of dams and road crossings have proved obstacles to the stream's natural flow, and hinders the ability for nutrients and organisms (including the migratory American Eel, *Anguilla Rostrata*, which lives in freshwater/estuaries and migrates ~3,000 miles to the Sargasso Sea to spawn) to travel upstream or downstream.

The Saw Mill River is a 23.5 mile tributary that drains a long, narrow basin in Westchester County stretching from Chappaqua to Yonkers, its course passes through residential, commercial, urban and forested areas within the 26.5 sq. mile watershed (Saw Mill River Coalition). The Saw Mill is a major natural resource in Westchester county and a critical riparian corridor, as it is the county's southernmost tributary to the Hudson River and provides some of the only remaining habitat in this densely populated region for wide variety of plants and animals (Saw Mill River Coalition). Both the river and its watershed have been extensively modified by urban development, transportation infrastructure, and flood control projects (Riverkeeper). The river's course is paralleled by the Saw Mill River Parkway for about 16 miles of the river's length (Saw Mill River Coalition). Construction of the parkway entailed moving a portion of the river channel, as well as co-construction of a trunk sewer line leading to Westchester County's largest wastewater treatment plant, in Yonkers (Riverkeeper). The river has suffered the impacts of flood control projects including straightening, relocating, and lining of the river channel;

filling lowlands; constructing flood walls; and replacing the river channel with a concrete flume (Riverkeeper; Groundwork Hudson Valley). As urban Yonkers grew, extensive sections of the Saw Mill River were completely covered. There is suspicion that untreated sewage is being released from the hidden trunk sewer line near Mills St. in the Getty Square area of urban Yonkers, due to abnormally elevated levels of fecal coliform detected at sites downstream (Guevara, 2016). In 2011, a decade's work by the Saw Mill River Coalition and Groundwork Hudson Valley culminated in the daylighting of a portion of the Saw Mill River in Yonkers, while additional areas have been daylighted since then (Riverkeeper). Daylighting is not far downstream of the suspected dumping site, while the site just above the suspected effluent zone, Walsh Rd., is accessible for sampling. Therefore, sampling at both sites could provide evidence that there is a source of untreated effluent being released into the stream under the cover of urban infrastructure (Guevara, 2016). Such illegal dumping practices further exacerbate the issue of water quality, and with an excessive concentration of metals (USGS, Wall et al 1998) as well as high levels of PCBs, orthophosphate, chlorodane and fecal coliform (Rogers, 1987; Philip, 1984), the Saw Mill is a main pollutant of the Hudson River and is a NYSDEC priority watershed (Groundwork Hudson Valley; Guevara, 2016).

Thus in observance of such differences among these tributaries to the mighty Hudson River, a burning question develops in how the differences of each landscape influence the ecology of each stream and, fundamentally, how the stream microbial community structures differ in each due to anthropogenic activity. I hypothesize that:

1. External inputs can change microbial communities in a detectable way.
   I test this by comparing microbial communities at an urban and a rural site before and after a known point source of treated and untreated sewage contamination.

2.0 Materials and Methods

2.1 Saw Kill Watershed Sampling Sites

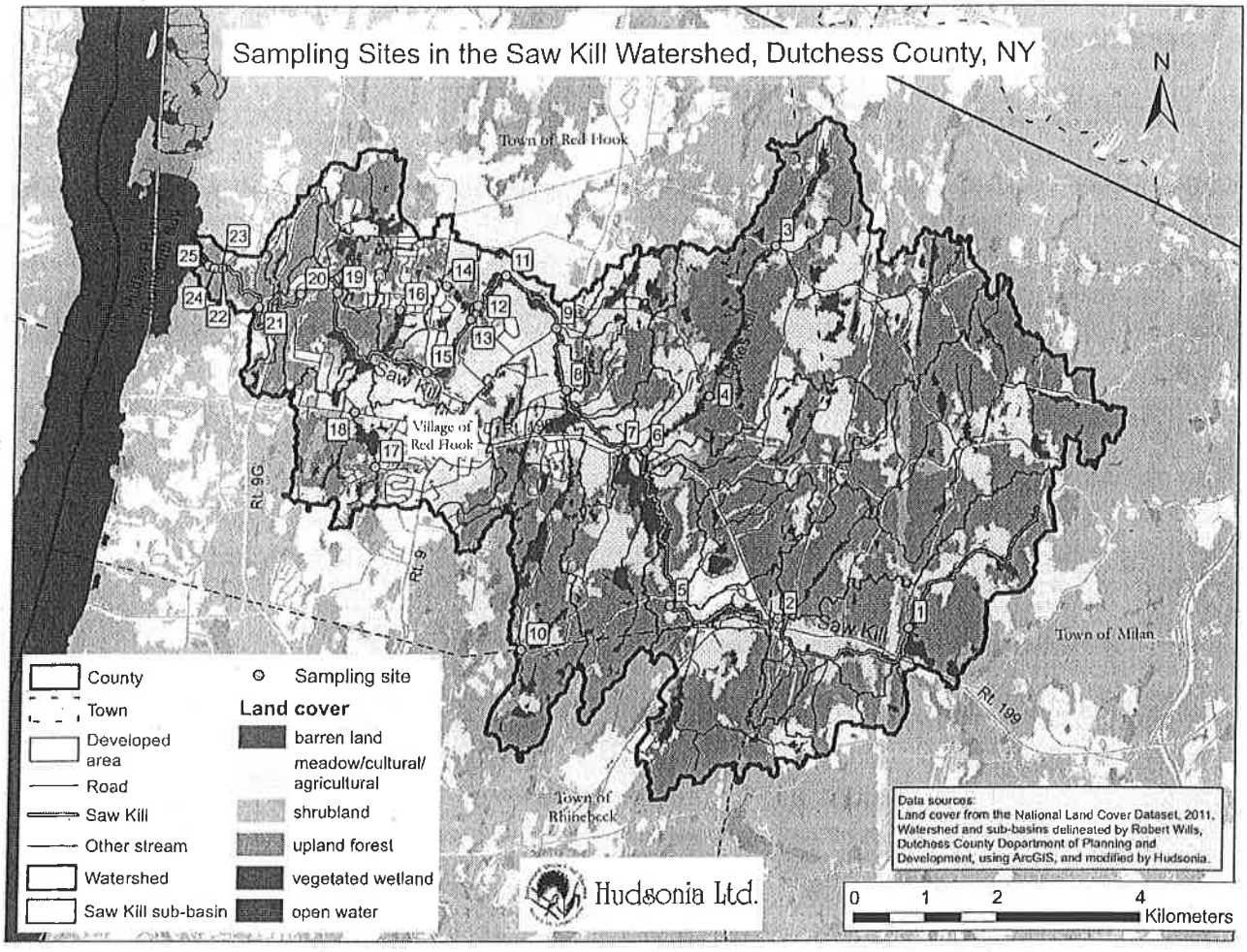


Figure 1. Saw Kill Watershed Sampling Sites. The focus of the study is closer to the mouth of the Saw Kill, at site 22 (below lower Saw Kill dam), site 23 (Bard WWTP effluent outflow), and site 24 (near the mouth of the tributary). Map created in ArcGIS by Chris Graham and Ben Houston to provide land cover analysis of the Saw Kill watershed for Marco Spodek's 2017 study "Nitrate Loading in the Saw Kill Watershed: small watershed nutrient dynamics, answering a community question, and assessing methodological approaches".

Sample sites Above Outflow (near and far pooled), Outflow, and Below Outflow (near and far pooled). Images from SKWC Flood Mitigation report

2.2 Saw Mill Watershed Sampling Sites

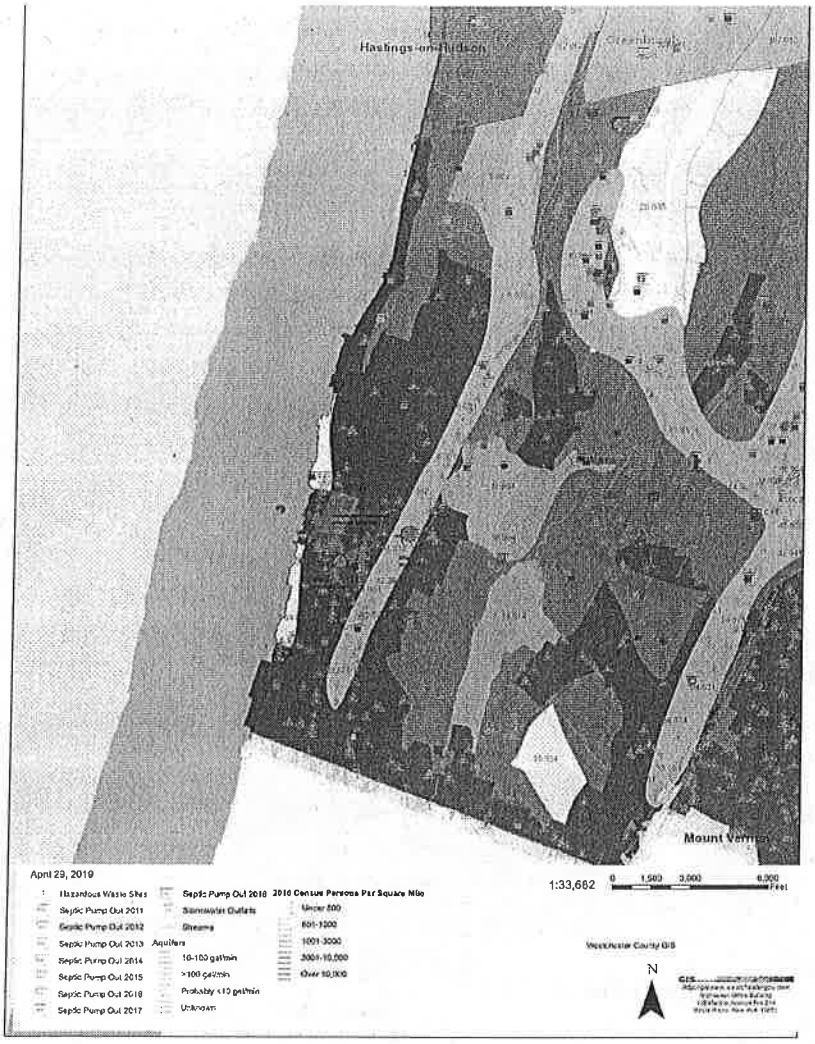
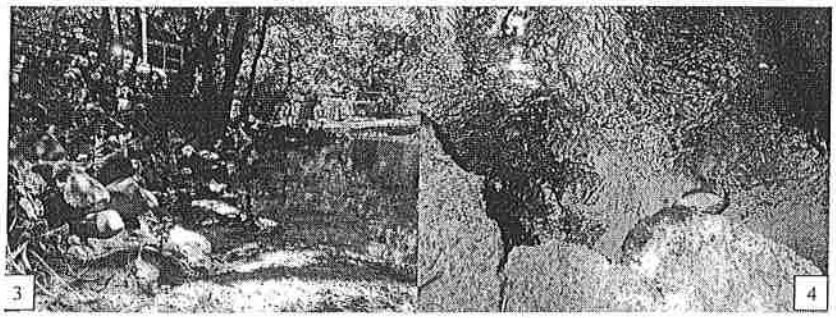


Figure 2. Saw Mill Watershed Sampling Sites. The focus of sampling efforts was at Walsh Rd., above the suspected untreated sewage effluent being released into the Saw Mill, and below the suspected effluent at the site of Daylighting, in Yonkers near Van Der Donck Park, and closer to the mouth of the tributary as it empties into the Hudson River. Map constructed within Westchester County Geographic Information Systems online database

Image C. Walsh Rd. sampling site along the Saw Mill River. [3] Shows Walsh Rd. and position in environment
[4] Shows where water samples were collected from
by Melissa Guevara (Guevara, 2016)



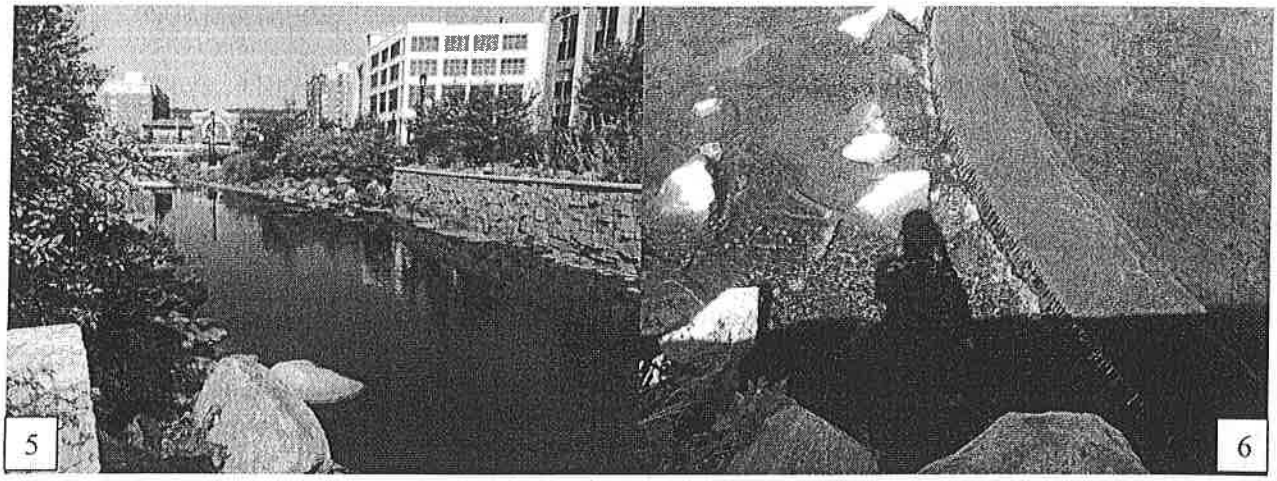


Image D. Daylighting phase along the Saw Mill River. [5] Shows how the Daylighting site and position in urbanized environment [6] Shows where water samples were collected from by Melissa Guevara (Guevara, 2016)

2.3 Field Sampling & Laboratory Analysis

No field sampling nor physical/chemical analyses were conducted within this study, as all sampling and processing of the Saw Kill was conducted by Bard students under the guidance of Prof. M. Elias Dueker and Prof. Gabriel Perron during the Bard Summer Research Institute (BSRI) in June and July of 2015, as well as September and October of 2015. The Saw Kill metadata[5] set, SK2015_MetaData, from 2015 includes water quality parameters such as total rainfall and rainfall 3 days prior to sampling, turbidity (NTU), sewage-associated bacteria counts (*E. coli, Coliform, Enterococcus*), dissolved oxygen (mg/L), percent dissolved oxygen, conductivity (ms/cm), salinity (ppt), and water temperature.

All sampling and physical/chemical analyses of the Saw Mill was conducted by former Bard student Melissa Guevara under the guidance of Prof. M. Elias Dueker over the months of August, September, and October 2015. The Saw Mill metadata set includes water quality

[5] Metadata - (also known as contextual data) refers directly to information regarding the original sample, the extraction and handling of the DNA, and the sequencing platform and data processing information (Field et al., 2011; Yilmaz et al., 2011)

parameters such as total rainfall and rainfall 3 days prior to sampling, turbidity (NTU), sewage-associated bacteria counts (*E. coli, Coliform, Enterococcus*), dissolved oxygen (mg/L), percent dissolved oxygen, conductivity (ms/cm), salinity (ppt), and water temperature.

Students collected water samples from the top 0.5 m of the water surface then placed in acid-washed and autoclaved 2L sample containers. Water samples were aseptically filtered through a 0.22 μm Sterivex filter, then filters were stored at -80°C. DNA was extracted from filters using the Qiagen/MoBio PowerWater extraction kits. To control for contamination, DNA was extracted from blank filters to serve as control. Amplicon pyrosequencing was then performed on the extracted DNA, with sequencing reactions prepared from each DNA extraction using the eubacterial forward primer 27F. DNA was then amplified through a single-step 30-cycle PCR machine to produce sequences.

2.4 Sequence Processing

All sequence processing was performed by Professor Dueker and Professor Perron. Sequence files were processed, aligned, and categorized independently using Illumina MiSeq platform with a custom Dada2 pipeline. In brief, raw sequence files were denoised using DADA2's default parameters (Pindling et al., 2018), then trimmed, requiring a sequence minimum length of 200 and allowing for 1 mismatch to the barcode and 2 mismatches to the primer. Then sequences were aligned against the SILVA ribosomal RNA gene database. After building ASV tables and chimeras were detected and removed, OTUs were assigned at the 97% identity threshold, using the 'assignTaxonomy' function of DADA2 against the SILVA ribosomal RNA gene database version 132 (Pindling et al., 2018). Since the study is focused on bacteria, all

sequences classified as "Chloroplast," Mitochondria,"and "Archaea" were removed from the bacterial dataset for downstream analyses.

2.5 Meta-analysis of Saw Kill and Saw Mill

In order to test the hypothesis of external inputs changing microbial communities in a detectable way, meta-analysis[6] was performed combining metagenomic data and metadata of a Suburban/Rural stream (Saw Kill) and an Urban stream (Saw Mill) known to be contaminated with treated and untreated sewage. Upon completion of the sequence processing, sequences were combined with environmental metadata during BSRI 2015 for the Saw Kill, and for analysis in phyloseq, a microbiome analysis package in R Core Team (2018). Rare OTUs were not removed from these analyses. All Saw Kill sediment samples were removed from the sample dataset into a subset, as this study only analyzes the aquatic microbial communities. Alpha and beta diversity statistics visualizations were acquired using phyloseq and ggplot2 (Wickham, 2009). Statistical tests of differences in diversity (using Shannon's H index) and similarity (using Bray-Curtis Dissimilarity) between habitats and sites were performed in the stats (R Core Team (2018)), ggpubr (Kassambara, 2018), and vegan (Oksanen et al., 2017) R packages (Dueker et al., 2018). ANOVA (statistical significance assigned at $p < 0.05$) and, if significant, Tukey *post hoc* (95% family-wise confidence level) tests were run on multiple-site comparisons to assess whether sites upstream and downstream of effluent outflow harbor significantly different microbial communities or have an affect on eachother. The composition of the bacterial communities across sites was compared by creating ordination plots that were designed using

[6] Meta-analysis - The process of performing comparative investigation of features between datasets, is greatly enhanced by the combination of metagenomic data and metadata (Knight et al., 2012)

non-metric dimensional scaling (nMDS) of a Bray-Curtis Dissimilarity matrix (Bray and Curtis, 1957; Qu et al., 2010; Rosi et al., 2018), and cluster ellipses were drawn in by hand to highlight sample clustering. Phylum-level relative abundance (phylum > 2%) of both streams was depicted using stacked bar plots in R package phyloseq. Genus-level relative abundance of the full Saw Kill microbial community was created by order of the highest OTU abundance of the 11 predominant bacterial genus' detected across Saw Kill sites, importing the list of ranked Genus-level relative abundance from the Saw Kill metagenomic dataset in R into a Microsoft Excel Spreadsheet. Cumulative OTU abundance and Genus-level abundance values were then calculated in Microsoft Excel to give visuals of aggregate sum OTU abundance and aggregate sum Genus abundance across Saw Kill sites. DESeq2 was used to identify over-represented taxonomic groups across both streams (Love et al., 2014).

2.5 Multivariate Analysis

In order to determine potential effects of parameters that can influence metagenomic data and meta-analysis, visual depictions of average turbidity (NTU), total rainfall (inches), rainfall (24, 48, & 72 hours prior to sampling), conductivity (ms/cm), and salinity (ppt) were created from each stream's metadata set in Microsoft Excel Spreadsheets. Such water quality parameters can provide supporting evidence to possible factors selectively driving each stream's microbial community composition, of both the local community at each sampling site and of the regional communities across sampling sites.

3.0 Results

3.1 Community Diversity Measures

A total of 372,644 sequences were yielded from Illumina MiSeq sequencing of 16s rRNA in water samples of each stream (289,789 sequences from Saw Kill and 103,285 sequences from the Saw Mill). OTU analysis indicated diverse microbial assemblages across sites of the Saw Kill, resulting in identification of 9,566 OTUs at the level of 97% identity. To compare overall diversity within water samples at each location, alpha diversity estimates for the full microbial community of the Saw Kill were measured by Observed index and Shannon H index at each site. There was greater diversity at the Outflow than Above Outflow and Below Outflow (**Figure 3**). Above Outflow had a significant effect on Outflow in both Observed and Shannon indices (Shannon adjusted p=0.0069, Observed adjusted p=0.00033), while the Outflow had a significant effect on Below Outflow (Shannon adjusted p=0.0023, Observed adjusted p=0.000067)(ANOVA, Tukey *post hoc*, $p < 0.05$). In contrast, there was no significant relationship between Above Outflow and Below Outflow, nor between the Extraction Control and any of the three Saw Kill sites, indicating there was little or no contamination during the extraction process.

Alpha diversity estimates for the full microbial community of the Saw Mill showed no significant difference in the diversity of microbial communities across sites Walsh Rd. and Daylighting, for both Observed and Shannon indices (**Figure 4**).

Community composition and Beta Diversity estimates were assessed for both the full Saw Kill microbial community and the full Saw Mill microbial community by overall community comparison across sites of each stream. Non-metric dimensional scaling (NMDS) of

a Bray-Curtis dissimilarity matrix was used to compare microbial communities by habitat across Saw Kill sites, and between Walsh Rd. and Daylighting. Bacterial composition comparison of Bard Outflow to Pooled Water sites (Above Outflow and Below Outflow) revealed significant differences in community composition between the pooled water sites and the Outflow. Cluster ellipses reveal significant differences in the microbial community composition between pooled water sites (Above and Below Outflow) and the Outflow, as Outflow community shows tighter clustering than the other two clustered sites (**Figure 5**). Beta Diversity estimates for Saw Mill sites at Walsh Rd. and Daylighting revealed similar community composition across sites, with cluster ellipses giving a strong confidence level of community similarity across sites **(Figure 6).**

3.2 Relative Abundance Comparison

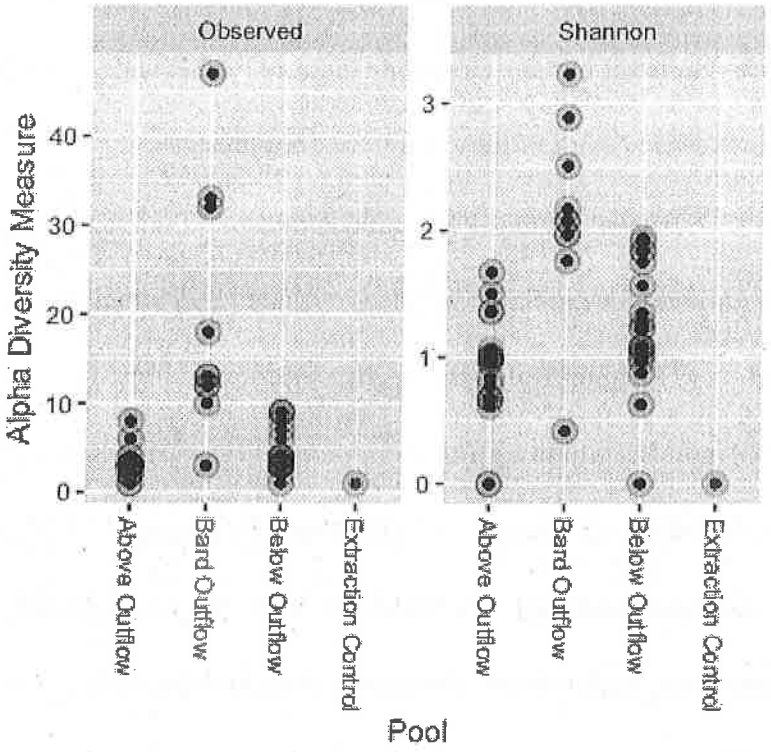
Relative abundance of taxa across the full Saw Kill microbial community and the full Saw Mill microbial community revealed the most dominant bacteria across sites and habitat. To avoid false reporting in meta-analysis of heterogeneity across stream microbial communities, in which microbial communities show greater than random spatial similarity no matter the method of measuring diversity, but the magnitude of heterogeneity detected was greater if a lower-resolution taxonomic definition was intentionally utilized (Horner-Devine et al., 2007; Zeglin, 2015). Thus, a qualitative approach was taken by classifying taxa at the more abundant Phylum-level as this is a broad classification of Bacteria, as well as classifying taxa at the Genus-level, giving a more specific and accurate understanding of the stream microbial compositions by abundance.

Relative Abundance at the Phylum-level (Phylum >2%) across Saw Kill sites (**Figure 7**) revealed *Proteobacteria* to be most abundant across sites (~48% Above Outflow, ~70% at

Outflow, and ~50% Below Outflow, as well as ~5% in the Extraction Control) followed by *Acidobacteria* (~30% Above Outflow, ~20% at Outflow, and ~25% Below Outflow). *Actinobacteria* appeared Above Outflow (~22%) and Below Outflow (~20%) more than at the Outflow (~8%), while *Firmicutes* was by dominant in the Extraction Control (~95%) and appeared in trace amount at the Outflow (~4%). Bard Outflow had trace abundance of *Planctomycetes* (~3%), *Nitrospirae* (~4%), *Verrumicrobia* (~2%), and *Chlamydiae* (~2%).

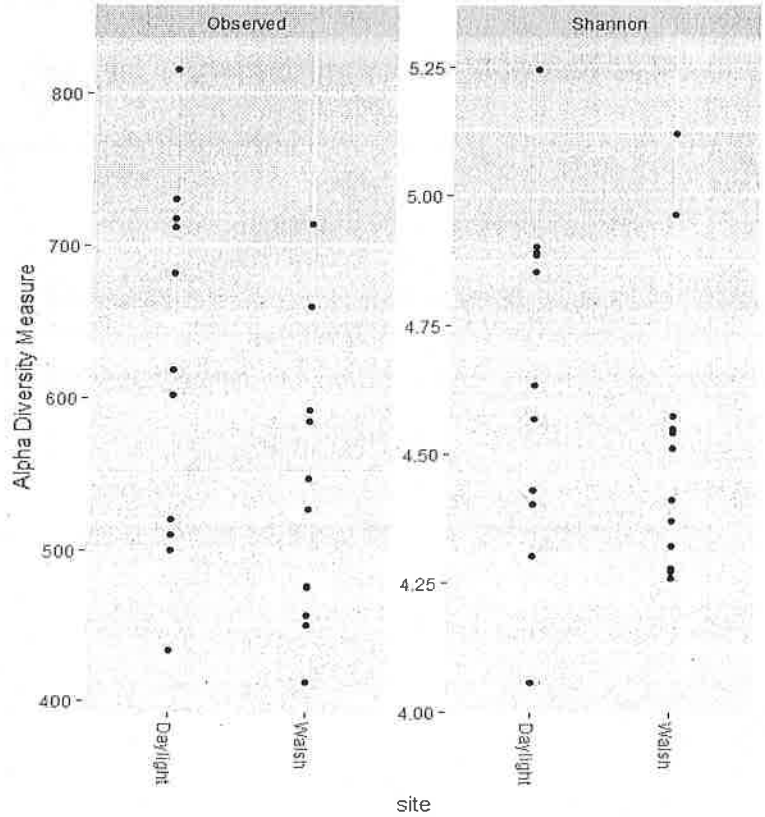
Relative Abundance at the Phylum-level for the Saw Mill also revealed *Proteobacteria* to be most abundant across sites Walsh Rd. (~47%) and Daylighting (~48%), followed by *Actinobacteria* (~28% for both Walsh Rd. and Daylighting), *Bacteroidetes* (~12% Walsh Rd. and ~13% Daylighting). *Cyanobacteria* (~2% Walsh Rd. and ~4% Daylighting), *Firmicutes* (~2 Walsh Rd. and ~3% Daylighting), and *Armatimonadetes* (~2% Walsh Rd. and ~3% Daylighting) were present in trace amounts across both sites, with relative abundance plots looking very similar for both Saw Mill sites (**Figure 8).**

DESeq2 (**Figure 9)** determined the top 39 overrepresented Genus according to the most abundant Phyla across both the Saw Kill and the Saw Mill. For a particular gene, a log2 fold change of -1 for a condition treated (Outflow) vs untreated (water) means the treatment induces a change in observed expression level of 2^-1 = 0.5 compared to the untreated condition (https://support.bioconductor.org/p/62927/). Thus for the Saw Kill, data points represented in positive territory display genus overrepresented in water, and negative data points show which genus are overrepresented in the Outflow.



Saw Kill

Figure 3. Alpha diversity estimates of Saw Kill sites Above Outflow, Outflow, and Below Outflow. Bacterial communities were more diverse at the Outflow and Below Outflow. Shannon's H index displays both evenness and abundance of the microbial community, showing a more even and slightly less clustered dispersal than the observed measure.

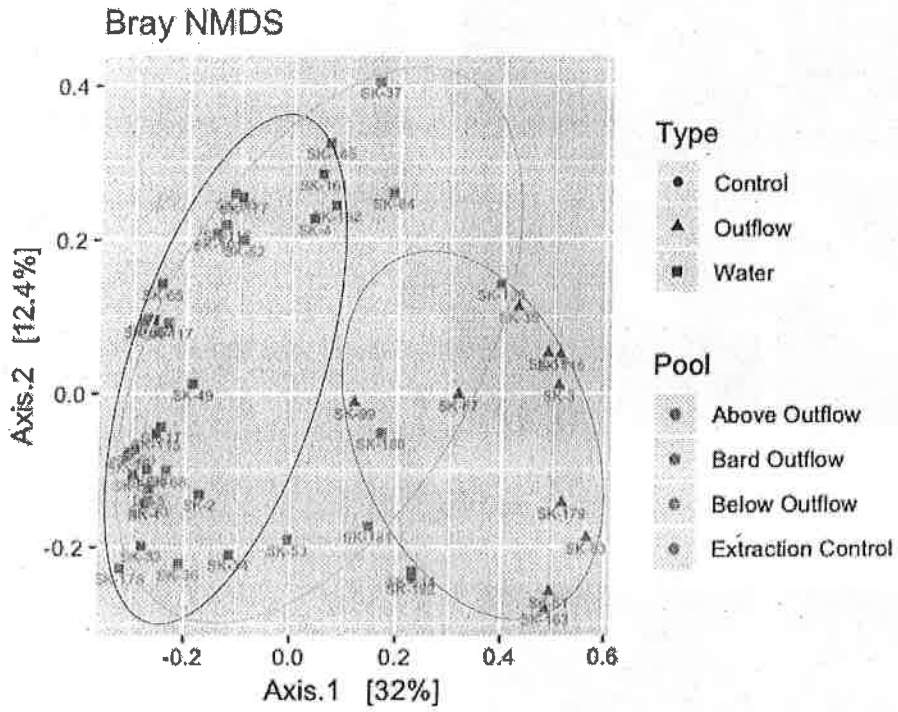


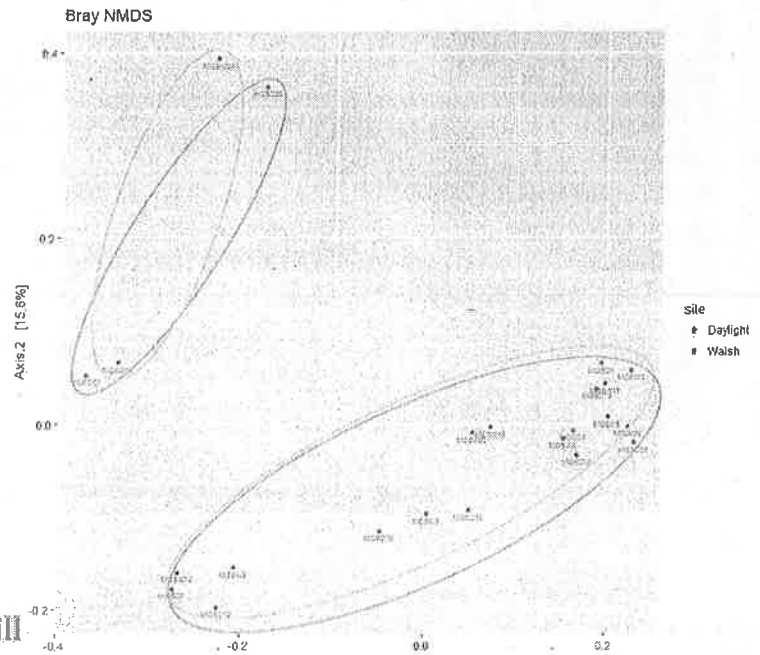
Saw Mill

Figure 4. Alpha Diversity estimates for Saw Mill sites Walsh Rd. and Daylighting. Bacterial communities were more diverse at Daylighting than Walsh Rd. Shannon H index shows the relative evenness and less dispersal at Walsh Rd. than Daylighting, as compared to the Observed reading which shows greater dispersal.

Saw Kill

Figure 5. Non-metric multidimensional scaling (NMDS) analysis of the dissimilarities among Saw Kill microbial communities using Bray-Curtis distances, at sites Above Outflow, Outflow, and Below Outflow. Bray-Curtis distances were calculated using the relative abundance of OTUs in each sample. Percentages in axes represent % variation explained by that axis. Cluster ellipses were drawn by hand to highlight sample clustering.

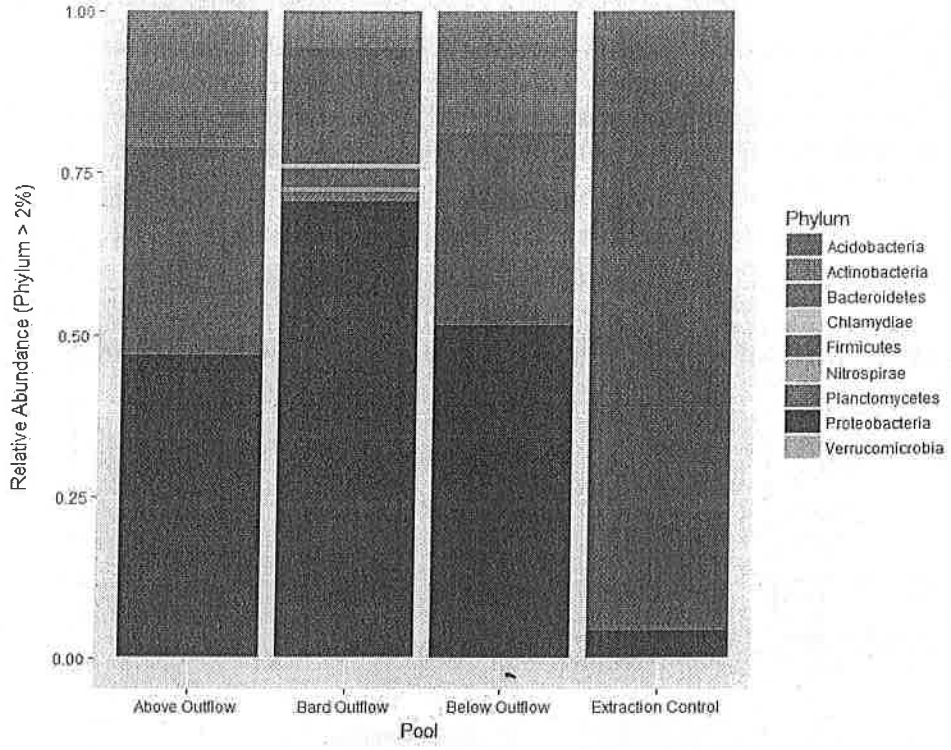




Saw Mill

Figure 6. Non-metric multidimensional scaling (NMDS) analysis of the dissimilarities among Saw Mill microbial communities, at sites Walsh Rd. and Daylighting. Bray-Curtis distances were calculated using the relative abundance of OTUs in each sample. Percentages in axes represent % variation explained by that axis. Cluster ellipses were drawn by hand to highlight sample clustering.

Figure 7. Phylum-level relative abundance across Saw Kill sites (Above Outflow, Outflow, Below Outflow, and Extraction Control). Only the phyla with a relative abundance >2% are shown across sites of each stream.



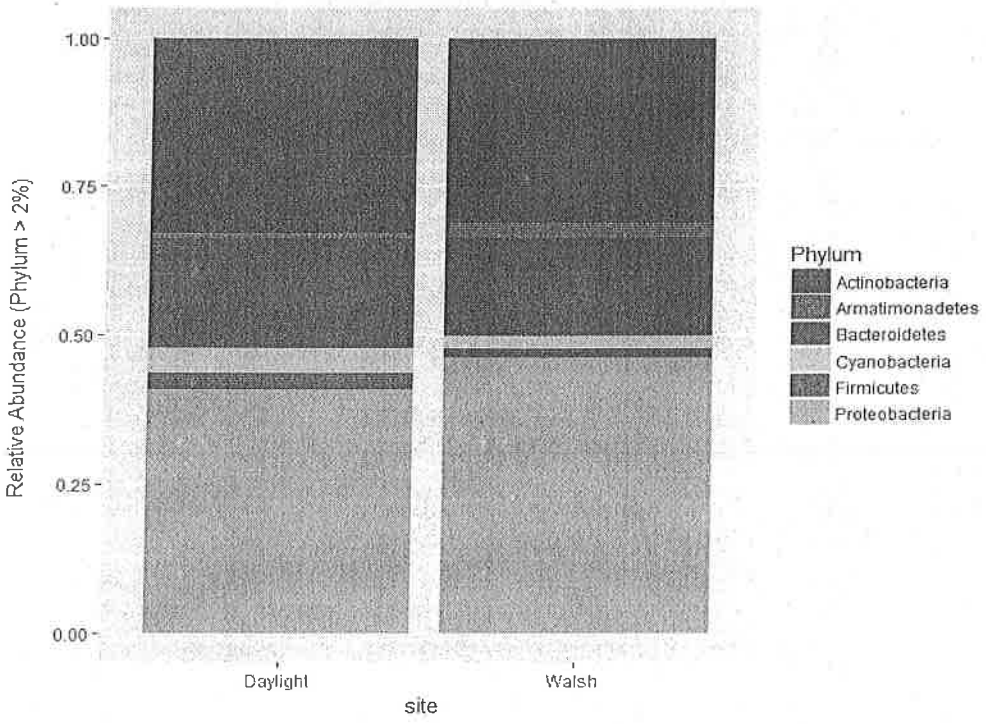


Figure 8. Phylum-level relative abundance across Saw Mill sites Walsh Rd. and Daylighting. Only the phyla with a relative abundance >2% are shown across sites of each stream.

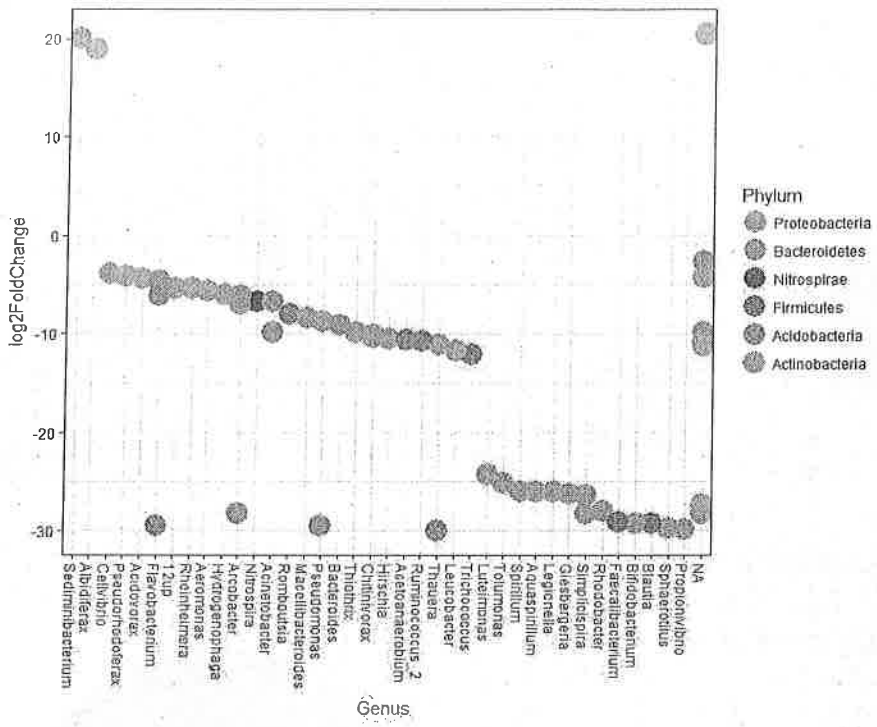


Figure 9. Microbial communities of all Saw Kill water sites (pooled Above Outflow and Below Outflow) and the Outflow. Data points in positive territory display genus overrepresented in water, and negative data points show which genus are overrepresented in the Outflow.

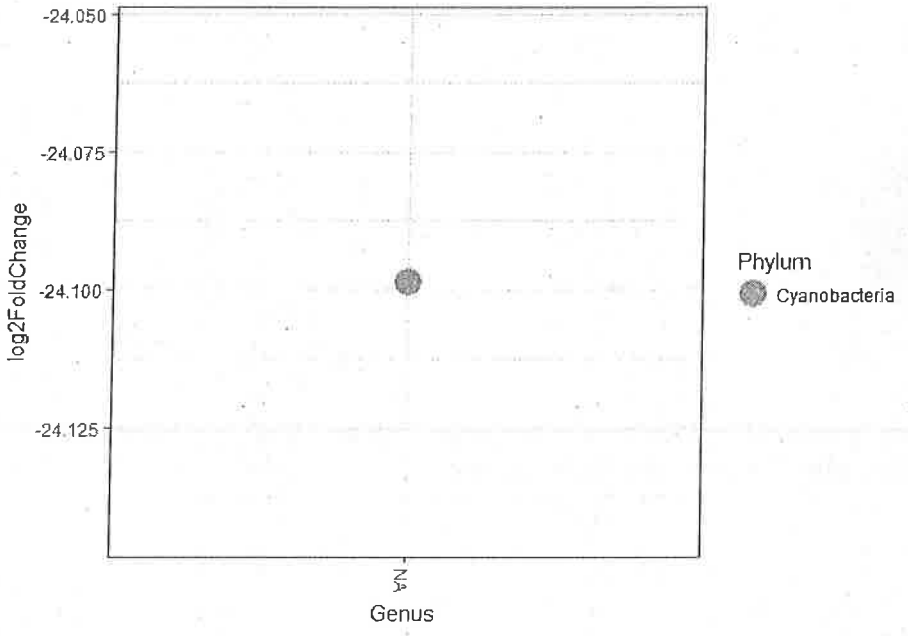
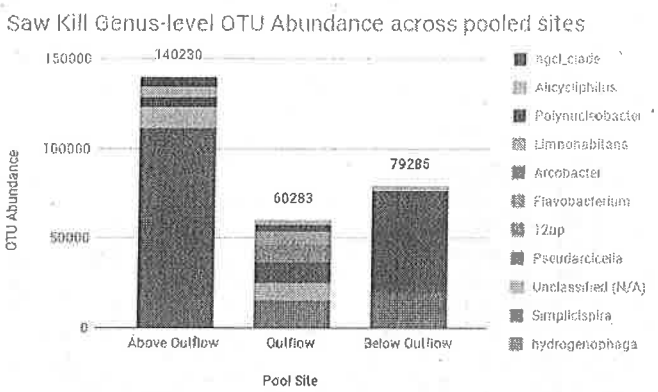


Figure 10. Microbial communities of Saw Mill sites Walsh Rd. and Daylighting. Only Cyanobacteria, at the phylum level, showed up as different, with presence of Cyanobacteria detected at Daylighting and not Walsh Rd.

11A



11B

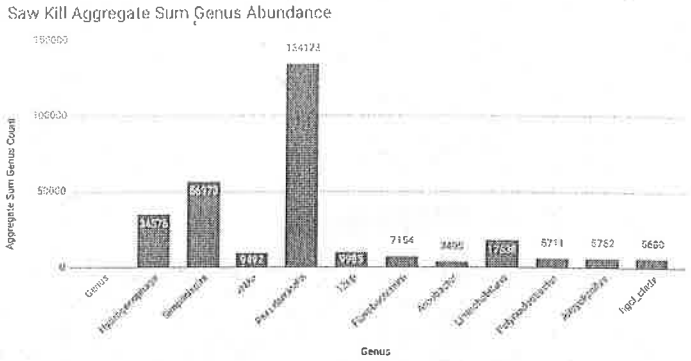
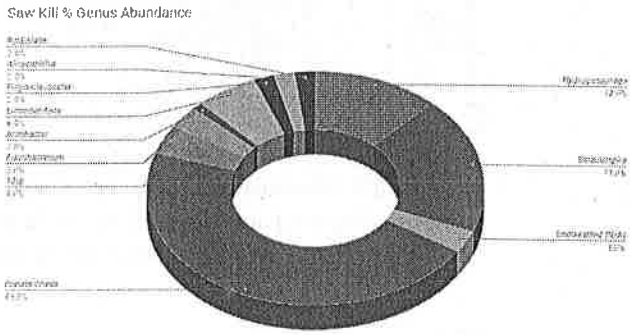


Figure 11.
(A) Saw Kill Genus-level OTU abundance and total abundance across pooled sites. **(B)** Saw Kill aggregate sum Genus-level abundance **(C)** Saw Kill % Genus-level abundance

11C



12A

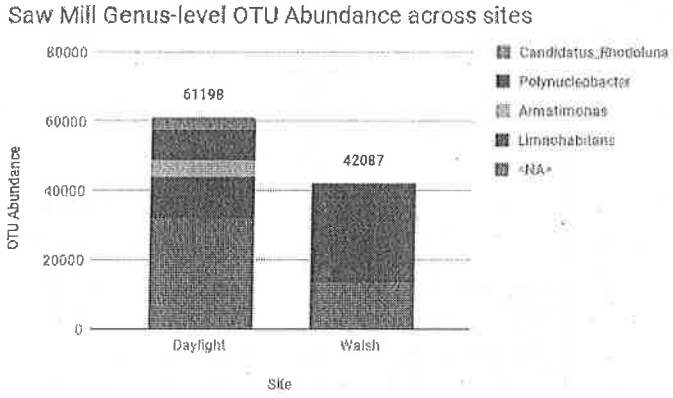
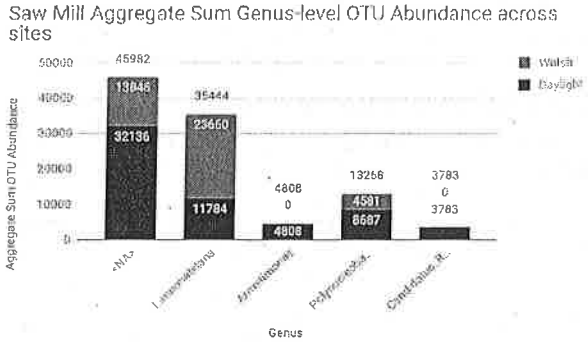
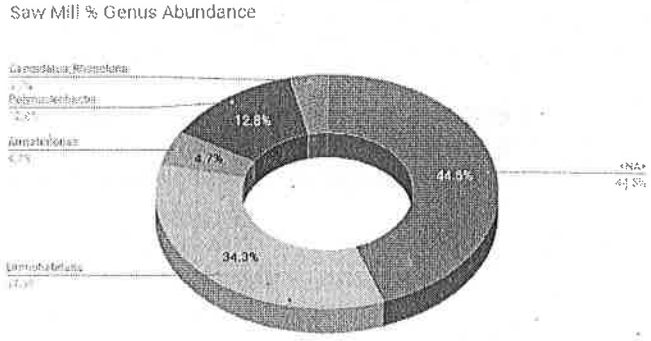


Figure 12. (A) Saw Mill Genus-level OTU abundance across sites **(B)** Saw Mill aggregate sum Genus-level abundance **(C)** Saw Mill % Genus-level abundance

12B



12C



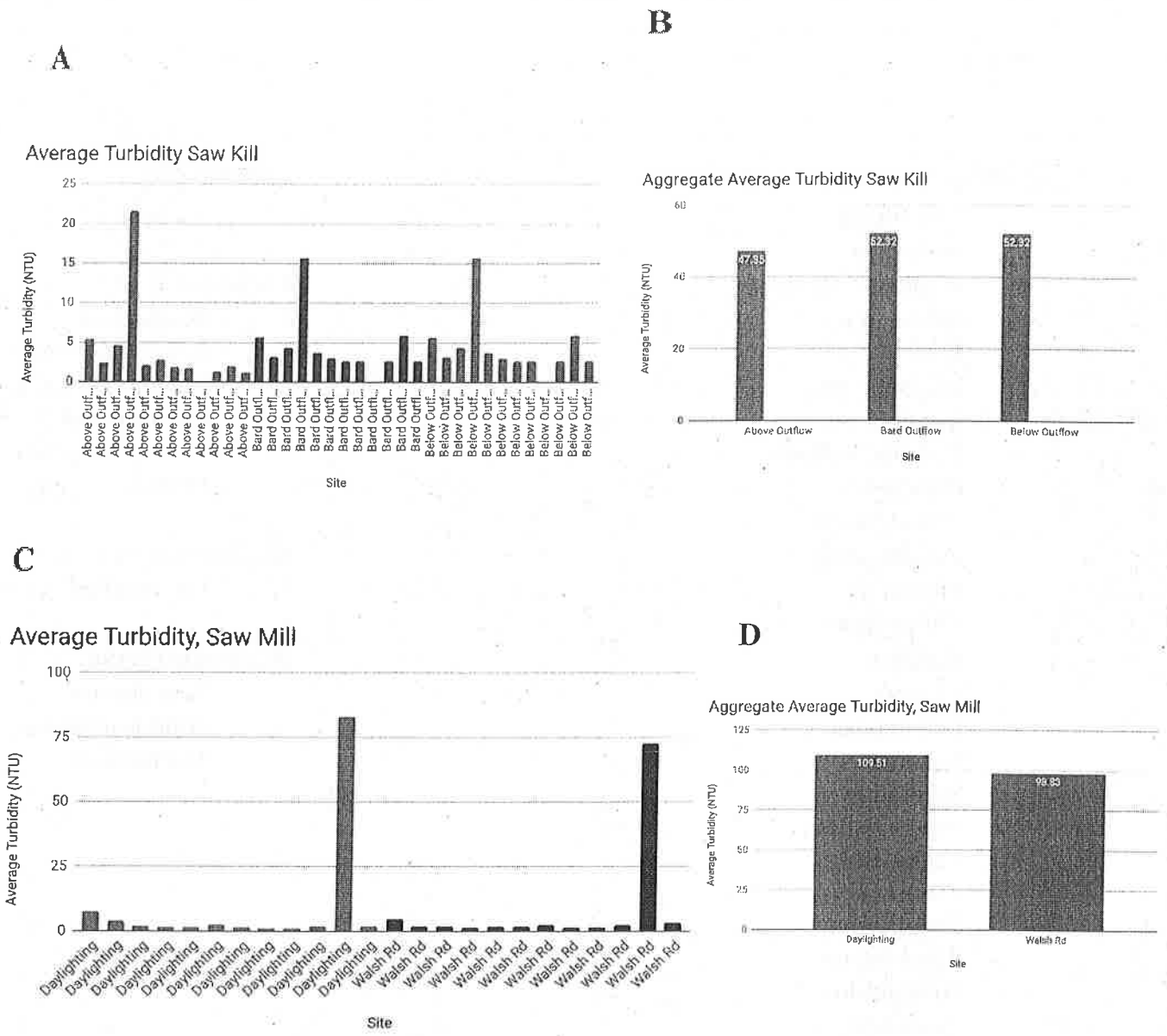


Figure 14. Average Turbidity of Saw Kill and Saw Mill
(A) Average Turbidity of Saw Kill at Above Outflow, Outflow, and Below Outflow over sampling days.
(B) Aggregate Turbidity of Saw Kill at each pooled site.
(C) Average Turbidity of Saw Mill at Walsh Rd. and Daylighting over sampling days.
(D) Aggregate Turbidity of Saw Mill at each site.

Table 1. Genus-level taxonomic assignment of OTUs to their respective Phylum, detected in samples across Saw Kill sites.

Phylum

Genus

Proteobacteria

- Albidiferax
- Cellvibrio
- Pseudorhodoferax
- Acidovorax
- 12up
- Rheinheimera
- Aeromonas
- Hydrogenophaga
- Arcobacter
- Acinetobacter
- Pseudomonas
- Thiothrix
- Chitinivorax
- Hirschia
- Thauera
- Luteimonas
- Tolumonas
- Spirillum
- Aquaspirillum
- Legionella
- Giesbergeria
- Simplicispira
- Rhodobacter
- Sphaerotilus
- Propionivibrio
- Unclassified

Bacteroidetes

- Sediminibacterium
- Flavobacterium
- Macellibacteroides
- Bacteroides
- Unclassified

Nitrospirae

- Nitrospira

Firmicutes

- Romboutsia
- Acetoanaerobium
- Ruminococcus_2
- Trichococcus
- Faecalibacterium
- Blautia

Acidobacteria

- Unclassified

Actinobacteria

- Leucobacter
- Bifidobacterium
- Unclassified

Table 2. Taxonomic assignment of OTUs at Phylum and Genus level, and their rank abundance at each Saw Kill pooled site on each sampling date.

OTU	Phylum	Genus	Abundance	Pool Site	Date
404681	Proteobacteria	Hydrogenophaga	19097	Below Outflow	7/13/2015
400801	Proteobacteria	Simplicispira	10818	Above Outflow	10/6/2015
400783	Proteobacteria	Simplicispira	10804	Above Outflow	10/6/2015
397431	Proteobacteria	<NA>	9492	Bard Outflow	10/20/2015
404669	Proteobacteria	Hydrogenophaga	7677	Bard Outflow	7/8/2015
317890	Bacteroidetes	Pseudarcicella	7561	Bard Outflow	7/13/2015
425167	Proteobacteria	12up	7342	Bard Outflow	10/6/2015
240242	Bacteroidetes	Flavobacterium	7154	Bard Outflow	10/6/2015
317877	Bacteroidetes	Pseudarcicella	6663	Above Outflow	7/6/2015
400791	Proteobacteria	Simplicispira	6663	Above Outflow	10/20/2015
317871	Bacteroidetes	Pseudarcicella	6601	Below Outflow	7/13/2015
317887	Bacteroidetes	Pseudarcicella	6188	Above Outflow	10/6/2015
317864	Bacteroidetes	Pseudarcicella	5872	Above Outflow	6/22/2015
317865	Bacteroidetes	Pseudarcicella	5477	Below Outflow	7/8/2015
317894	Bacteroidetes	Pseudarcicella	5194	Below Outflow	7/6/2015
317879	Bacteroidetes	Pseudarcicella	5073	Above Outflow	7/15/2015
317883	Bacteroidetes	Pseudarcicella	5047	Above Outflow	7/8/2015
400777	Proteobacteria	Simplicispira	5005	Above Outflow	10/20/2015
317906	Bacteroidetes	Pseudarcicella	4947	Below Outflow	6/22/2015
317873	Bacteroidetes	Pseudarcicella	4920	Above Outflow	7/8/2015
317893	Bacteroidetes	Pseudarcicella	4755	Above Outflow	10/6/2015
317885	Bacteroidetes	Pseudarcicella	4648	Below Outflow	7/8/2015
317866	Bacteroidetes	Pseudarcicella	4634	Above Outflow	6/29/2015
317874	Bacteroidetes	Pseudarcicella	4499	Above Outflow	6/29/2015
317872	Bacteroidetes	Pseudarcicella	4308	Above Outflow	6/24/2015
400793	Proteobacteria	Simplicispira	4289	Below Outflow	10/20/2015
317908	Bacteroidetes	Pseudarcicella	4209	Bard Outflow	6/29/2015
404680	Proteobacteria	Hydrogenophaga	4126	Bard Outflow	7/15/2015
317901	Bacteroidetes	Pseudarcicella	3956	Above Outflow	6/22/2015
317895	Bacteroidetes	Pseudarcicella	3923	Below Outflow	7/6/2015
400799	Proteobacteria	Simplicispira	3791	Below Outflow	10/20/2015
404647	Proteobacteria	Hydrogenophaga	3676	Bard Outflow	6/24/2015
317880	Bacteroidetes	Pseudarcicella	3627	Above Outflow	7/13/2015
317911	Bacteroidetes	Pseudarcicella	3593	Above Outflow	7/13/2015
317881	Bacteroidetes	Pseudarcicella	3580	Below Outflow	6/24/2015
317870	Bacteroidetes	Pseudarcicella	3552	Above Outflow	7/6/2015
400796	Proteobacteria	Simplicispira	3495	Above Outflow	6/22/2015
290037	Proteobacteria	Arcobacter	3405	Bard Outflow	7/1/2015
400808	Proteobacteria	Simplicispira	3382	Below Outflow	6/22/2015
317869	Bacteroidetes	Pseudarcicella	3348	Below Outflow	6/22/2015
317912	Bacteroidetes	Pseudarcicella	3303	Below Outflow	6/29/2015
317903	Bacteroidetes	Pseudarcicella	3281	Below Outflow	6/24/2015
401166	Proteobacteria	Limnohabitans	3137	Above Outflow	6/22/2015
401180	Proteobacteria	Limnohabitans	3052	Below Outflow	7/13/2015
317884	Bacteroidetes	Pseudarcicella	2998	Below Outflow	6/29/2015
401189	Proteobacteria	Limnohabitans	2998	Bard Outflow	7/13/2015
422693	Proteobacteria	Polynucleobacter	2980	Above Outflow	10/6/2015
401201	Proteobacteria	Limnohabitans	2930	Above Outflow	7/15/2015
402468	Proteobacteria	Alicycliphilus	2881	Above Outflow	10/6/2015

152608	Actinobacteria	hgcI_clade	2874	Above Outflow	10/6/2015
402446	Proteobacteria	Alicycliphilus	2871	Above Outflow	10/6/2015
317876	Bacteroidetes	Pseudarcicella	2826	Below Outflow	7/15/2015
317891	Bacteroidetes	Pseudarcicella	2819	Above Outflow	6/24/2015
400776	Proteobacteria	Simplicispira	2818	Above Outflow	7/6/2015
152616	Actinobacteria	hgcI_clade	2806	Above Outflow	10/6/2015
401175	Proteobacteria	Limnohabitans	2734	Above Outflow	10/6/2015
422703	Bacteroidetes	Polynucleobacter	2731	Above Outflow	10/6/2015
317898	Proteobacteria	Pseudarcicella	2721	Below Outflow	7/15/2015
401188	Proteobacteria	Limnohabitans	2687	Above Outflow	10/6/2015
400788	Proteobacteria	Simplicispira	2686	Above Outflow	6/22/2015
425146	Proteobacteria	12up	2643	Bard Outflow	6/22/2015
400775	Proteobacteria	Simplicispira	2622	Above Outflow	6/24/2015

Table 3. Taxonomic assignment of OTUs at Phylum and Genus level, and their order of highest rank abundance at each Saw Mill site on each sampling date.

OTU	Phylum	Genus	Abundance	Site	Date
89198	Cyanobacteria	<NA>	8973	Daylight	9/7/2015
23122	Proteobacteria	Limnohabitans	8326	Walsh	10/5/201
53028	Actinobacteria	<NA>	7646	Daylight	8/31/201
53026	Actinobacteria	<NA>	6373	Daylight	9/2/2015
53020	Actinobacteria	<NA>	5170	Walsh	8/31/201
53016	Actinobacteria	<NA>	4838	Daylight	9/7/2015
55426	Armatimonadetes	Armatimonas	4808	Walsh	9/23/201
31746	Proteobacteria	Polynucleobacter	4795	Daylight	9/16/201
53035	Actinobacteria	<NA>	4599	Walsh	9/7/2015
31763	Proteobacteria	Polynucleobacter	4581	Walsh	9/16/2015
53018	Actinobacteria	<NA>	4306	Daylight	9/9/2015
23132	Proteobacteria	Limnohabitans	4242	Daylight	10/5/2015
24372	Proteobacteria	Limnohabitans	4162	Walsh	8/31/2015
53029	Actinobacteria	<NA>	4077	Walsh	9/2/2015
31749	Proteobacteria	Polynucleobacter	3892	Daylight	9/14/2015
24380	Proteobacteria	Limnohabitans	3850	Daylight	8/25/2015
24362	Proteobacteria	Limnohabitans	3835	Walsh	8/25/2015
49711	Actinobacteria	Candidatus_Rhodoluna	3783	Daylight	9/16/2015
24379	Proteobacteria	Limnohabitans	3748	Walsh	9/21/2015
24374	Proteobacteria	Limnohabitans	3692	Daylight	9/2/2015
24373	Proteobacteria	Limnohabitans	3589	Walsh	9/7/2015

4.0 Discussion

The study provides a unique insight into the comparison of microbial community diversity across a Suburban/Rural stream (Saw Kill) and an Urban stream (Saw Mill), as there have been relatively few studies analyzing the changing aquatic microbial communities in the face of increasing urbanization and population density. Such differences in diversity can be attributed to local sources, as well as being selectively driven by environmental heterogeneity and climate change, thus it is not entirely surprising that there is higher diversity in observed Saw Kill microbial communities than Saw Mill microbial communities. Across both stream gradients, an external input in the form of treated and untreated sewage is released into the stream, evidently altering the microbial communities at the sites of outflow and at sites immediately downstream.

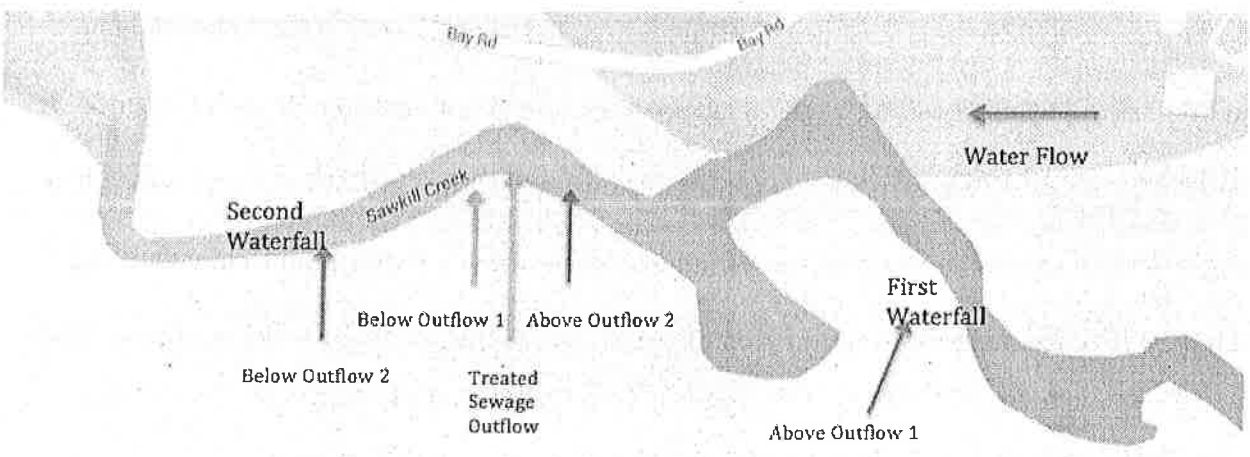
Bacterial communities across sites along the Saw Kill showed relatively similar diversity measures, with slightly more diverse communities at the Outflow and Below Outflow than Above Outflow, for the full Saw Kill microbial community (**Figure 3**). Higher diversity measures at the Outflow and Below Outflow can likely be attributed to exposure of novel bacteria from the human gut microbiome, as a diverse bacterial culture is introduced into the waterway through the treated wastewater effluent. As these bacteria are unlikely to originate from non-anthropogenic sources, bacteria associated with the gut microbiome are likely to influence the higher diversity measures at the Outflow and Below Outflow. Additionally, there appears greater evenness in the Shannon H index at sites Above Outflow and Below Outflow, which may be attributed to the introduction of novel bacteria through the effluent outflow. There

are roughly 10^30 microbiota living within our gut microbiomes, thus we share a functional core microbiome, but not a core microbiota as we humans are diverse in what we consume and how our microbiome respond to external, edible inputs. Bacterial cultures in Bard's WTTP rotating biological contactors (RBCs) could provide a source for novel sewage-associated bacteria as the turbines provide a host culture for microbiota to evolve and adapt to their new environment..

For future direction, consideration for future meta-analyses that combines multivariate data and metadata, is not always valid to compare the values of derived diversity metrics or the abundance of microorganisms based on data collected using different methodologies and taxonomic resolutions, so a fully quantitative meta-analysis, using a response index, was not possible. However, it is valid to accept significant results of a study as informative, no matter the data type. For example, a meta-analysis of heterogeneity in soil microbial communities showed greater-than-random spatial similarity no matter the technique used to measure diversity, but the magnitude of heterogeneity detected was greater if a lower-resolution taxonomic definition was utilized as to a more specific taxa classification that yields less results. Thus, using a range of taxonomic

5.0 List of Figures and Tables

6.0 Appendix



Samples were taken by students during BSRI 2015 from five sites along the Sawkill, a 14 mile tributary that flows westward into the Hudson River. Two sites above the outflow (near and far), the actual site of the outflow, and two sites below the outflow (near and far). A total of five water samples were taken on each sampling date, one at each site.

Sampling dates were: 6/22/15, 6/24/15, 6/29/15, 7/1/15, 7/6/15, 7/8/15, 7/13/15, 7/15/15, 9/22/15, 9/29/15, 10/6/15, and 10/20/15.

We began at the far below outflow site in order to avoid contamination and worked our way upstream. All sample collections were done facing the current with the collection vehicle or tool out in front, so that the sample site was not disturbed or contaminated. To collect water samples, we used a 2L bottle (prior to collecting these samples the 2L bottles were acid washed and autoclaved to ensure sterility and avoid contamination), washed out three times using the water from the collection site, and then collected the sample by setting the bottle into the top 0.5m of the water surface and letting the current run into it. Directly after collecting the sample, we placed it into an insulated, dark backpack containing ice packs in order to keep the bacteria from replicating at an unusual rate from new exposure to heat and light (this would skew the sample). Upon returning to the lab, we split the samples in order to perform the various assays.

2.3 Laboratory Analysis

Next to a flame and on a bench cleaned with Ethanol, the BSRI water quality group carefully poured out 500 mL of the water sample to be used for their assays.

Approximately 500 mL of sample water was sent through a .22 µl Sterivex filter attached to a peristaltic pump in order to filter out any larger particles that might inhibit the DNA extraction. Sometimes less water than 500 mL was sent through the pump because it would get clogged up. Autoclaved deionized water was sent through the pump between each sample to clean it. The Tygon pump tubes themselves were autoclaved between each sampling day as well. The Sterivex filters for each sample were then stored at -80°C. DNA was then extracted from filters using the Qiagen/MoBio PowerWater extraction kits.

Water DNA Extraction using the Qiagen/MoBio PowerWater DNA Isolation kit:

Per DNA sample, you need:

- 5 2ml collection tubes
- 1 2ml spin filter
- 1 microbead tube
- 1000µL PW1
- 200µL PW2
- 650µL PW3
- 650µL PW4
- 650µL PW5
- 70µL Nuclease free water

*Warm Solution PW1 prior to use at 55°C for 5-10 minutes. Use Solution PW1 while still warm. Check Solution PW3 and warm at 55°C for 5-10 minutes if necessary. Solution PW3 can be used while still warm.

1. Filter water samples using a reusable or disposable filter funnel attached to a vacuum source. Disposable filter funnels, containing 0.22 µm or 0.45 µm filter membranes, can be ordered from MO BIO Laboratories (see page 3). The volume of water filtered will depend on the microbial load and turbidity of the water sample. (Please see Types of Water Samples in the Hints and Troubleshooting Guide section of the Instruction Manual).
2. If using a reusable filter funnel, remove the upper portion of the apparatus. If using a MO BIO Laboratories filter funnel, remove the 100 ml upper portion of the filter cup from the catch reservoir by snapping it off.
3. Using two sets of sterile forceps, pick up the white filter membrane at opposite edges and roll the filter into a cylinder with the top side facing inward. Note: Do not tightly roll or fold the filter membrane.
4. Insert the filter into the 5 ml PowerWater® Bead Tube.
5. Add 1 ml of Solution PW1 to the PowerWater® Bead Tube. Note: Solution PW1 must be warmed to dissolve precipitates prior to use. Solution PW1 should be used while still warm. For samples containing organisms that are difficult to lyse (fungi, algae) an additional heating step can be included. See Alternate Lysis Method in the Hints and Troubleshooting Guide.
6. Secure the PowerWater® Bead Tube horizontally to a MO BIO Vortex Adapter, catalog number 13000-V1-15 or 13000-V1-5.
7. Vortex at maximum speed for 5 minutes.
8. Centrifuge the tubes ≤ 4000 x g for 1 minute at room temperature. The speed will depend on the capability of your centrifuge. (This step is optional if a centrifuge with a 15 ml tube rotor is not available, but will result in minor loss of supernatant).
9. Transfer all the supernatant to a clean 2 ml Collection Tube (provided). Draw up the supernatant using a 1 ml pipette tip by placing it down into the beads. Note: Placing the pipette tip down into the beads is required. Pipette more than once to ensure removal of all supernatant. Any carryover of beads will not affect subsequent steps. Expect to recover between 600-650 µl of supernatant depending on the type of filter membrane used.
10. Centrifuge at 13,000 x g for 1 minute.
11. Avoiding the pellet, transfer the supernatant to a clean 2 ml Collection Tube (provided).
12. Add 200 µl of Solution PW2 and vortex briefly to mix. Incubate at 4°C for 5 minutes.
13. Centrifuge the tubes at 13,000 x g for 1 minute.
14. Avoiding the pellet, transfer the supernatant to a clean 2 ml Collection Tube (provided).
15. Add 650 µl of Solution PW3 and vortex briefly to mix. Note: Check Solution PW3 for precipitation prior to use. Warm if necessary. Solution PW3 can be used while still warm.
16. Load 650 µl of supernatant onto a Spin Filter and centrifuge at 13,000 x g for 1 minute. Discard the flow through and repeat until all the supernatant has been loaded onto the Spin Filter. Note: A total of two loads for each sample processed are required.
17. Place the Spin Filter basket into a clean 2 ml Collection Tube (provided).

18. Shake to mix Solution PW4 before use. Add 650 µl of Solution PW4 and centrifuge at 13,000 x g for 1 minute.
19. Discard the flow through and add 650 µl of Solution PW5 and centrifuge at 13,000 x g for 1 minute.
20. Discard the flow through and centrifuge again at 13,000 x g for 2 minutes to remove residual wash.
21. Place the Spin Filter basket into a clean 2 ml Collection Tube (provided).
22. Add 50 µl of sterile HyClone water to the center of the white filter membrane.
23. Centrifuge at 13,000 x g for 1 minute.
24. Discard the Spin Filter basket. The DNA is now ready for any downstream application. No further steps are required.

Free the DNA to store (-20°C to -80°C)

Dylan Dahan's PCR methods

Amplicon pyrosequencing was then performed on the extracted DNA using standard Polymerase Chain Reaction (PCR) machine, where the quantity of DNA product is detected at the end of the reaction and amplified.

PCR:

Materials:

-3 ng/µl concentration of DNA samples
-PowerUp Sybrr Green Mastermix
-Bovine Serum Albumium (10mg/ml)
-Nuclease free water
-Forward and reverse primers (16s and int1)
-1.5/2 mL centrifuge tube
-One clear qPCR 96-well plate and one clear qPCR 96-well plate tape
-a qPCR thermocycler

Method:

-Excel was used to calculate the respective quantities of reagents based on the total number of DNA samples, as well as create a map of the DNA samples used.

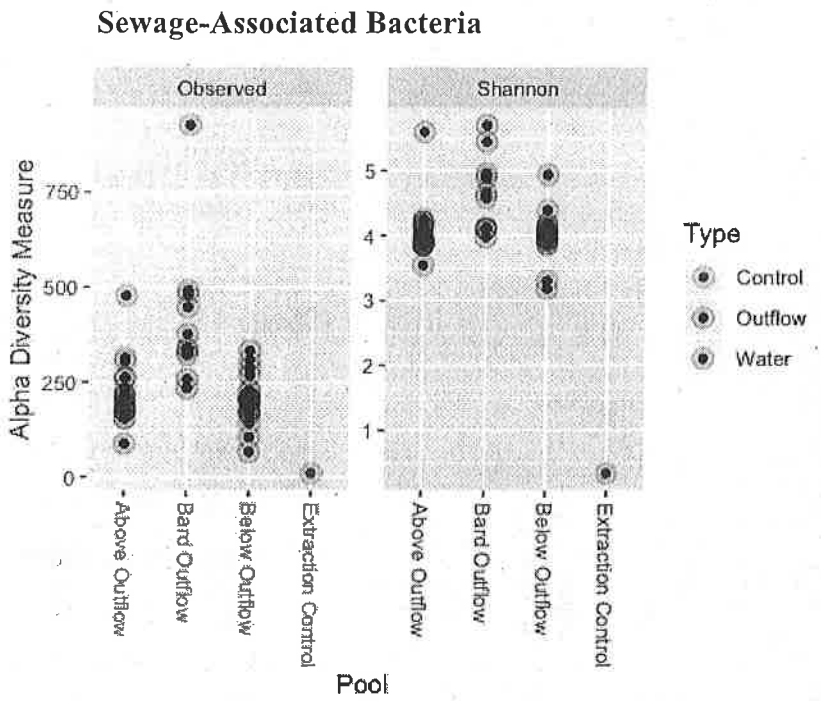
For one DNA sample:

Reagent	Volume Added (µl)
PowerUp Sybry Green MM	10
Bovine Serum Albumium (10 mg/ml)	0.2
Forward Primer	0.9
Reverse Primer	0.9
H_2O	6.0

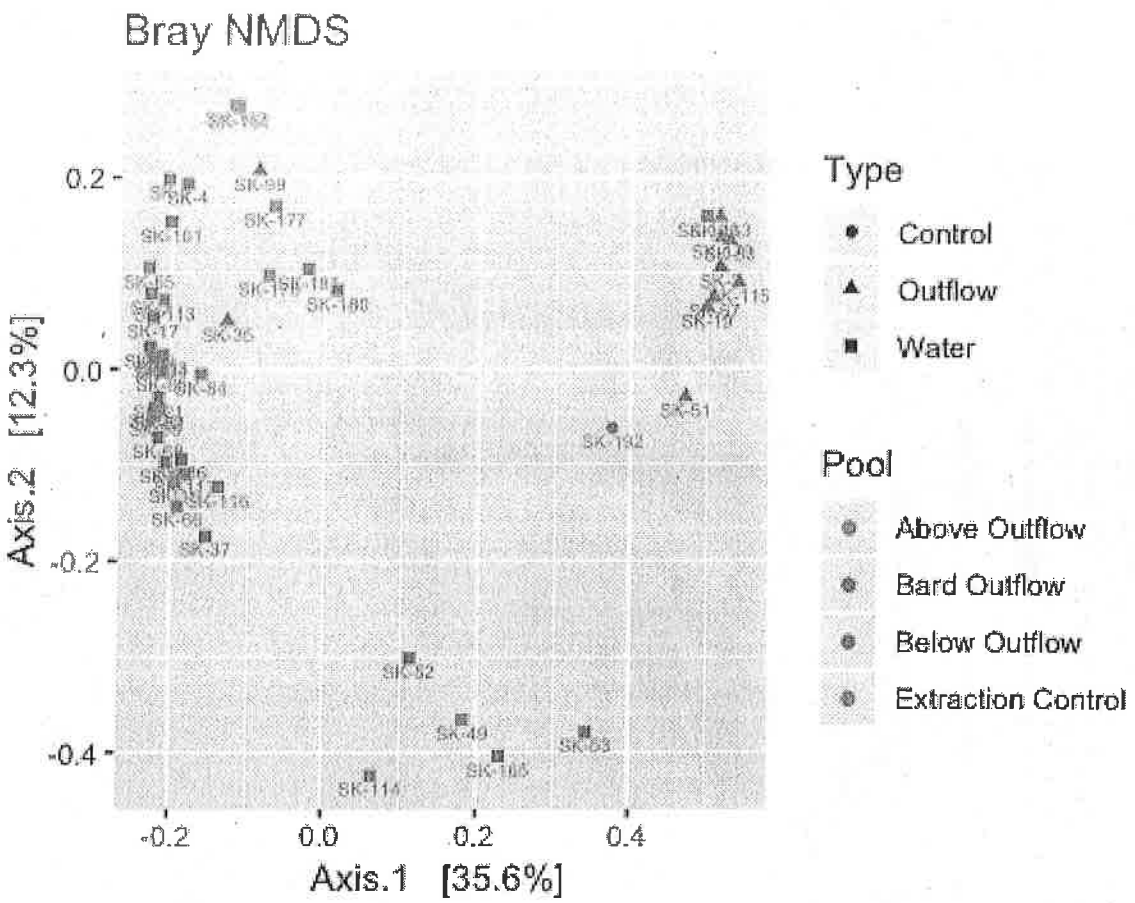
-Clean the whole lab bench thoroughly with a series of three chemical: Alcanox, 3% Hydrogen Peroxide, and 70% Ethanol. This includes bottoms of containers, pipettes, etc.
-Light a Bunsen burner to work under.
-Take out the materials from their respective locations and begin to thaw them. There should not be any ice crystals when pipetting.

-Combine the proper amount of each reagent in either a 1.5 or 2 ml centrifuge tube (depending on the total volume). Each reagent should be thawed and vortexed on medium-low three consecutive times before added to the Master Mix.
-After mix is assembled, take out a 96 well plate cover with the sticker in between pipetting.
-Use the same vortex technique and use the tabletop spinner for 3 seconds. Pipette 18 µl of the master mix in each reaction well. The Master Mix should be vortexed approximately every two rows to keep mixture homogenized. Change tips after each well and only pipette down to the first stop to avoid bubbles.
-Pipette 2 µl of each DNA sample in each well for a total volume of 20 µl. Make sure to use the same vortex technique for each sample. Each reaction will be done in triplicates; so 2 µl of each sample will go in to three separate wells. For NTCs use 2 µl nuclease free water.
-Seal the plate with clear 96-well plate qPCR tape and centrifuge on 300 rpm for 2 minutes to bring down any liquid on the sides.
-Set the thermocycler to:
1. 50°C for 2:00 min
2. 95°C for 10:00 min
3. 95°C for 0:20s
4. 60°C for 1:00 min
5. GO TO 3, 40 more times
6. Melt curve 65°C to 95°C, increment .5°C, for 0:05s
-Design the reaction plate on the qPCR to match the reaction plate created in the first step and make sure that SYBR fluorophores are selected.
Place the 96 well plate inside and wait! It runs for approximately 2 hours.
Finally, record the amplification curve, melt curve and melt peak and save to excel spreadsheet.

Further Supplementary material and supporting evidence by Sewage-Associated Bacteria Community Composition



Supplementary Figure 1. Alpha diversity estimates for the Saw Kill sewage microbial community. Bacterial communities were more diverse at the Outflow, and slightly more diverse Above Outflow, though not significantly different. Shannon's H index clustered the data closer together and absorbed an outlier at the Outflow, while the Observed measure is slightly more evenly dispersed, particularly Below Outflow.



Supplementary Figure 2. Beta diversity estimates showing differences of sewage microbial communities in relation to (NMDS) of Saw Kill sewage microbial community Above Outflow, Outflow, and Below Outflow. Percentages in axes represent % variation explained by that axis. Ellipses calculated using Euclidean distance (ggplot2, R package).

Sewage microbial communities across sites along the Saw Kill showed similar diversity measures to the full microbial community of the Saw Kill, albeit with slightly more diverse communities Above Outflow than Below Outflow, and highest diversity at the Outflow. This is consistent with the notion that sewage bacteria from a variety of peoples' gut microbiomes and sewage are released into the Saw Kill, increasing the abundance of sewage-associated bacteria. A slightly higher sewage bacteria diversity measure for Above Outflow could be attributed to sources of fecal contamination upstream in the Village of Red Hook and effluent pipes from septic systems leaking sewage into the Saw Kill.

7.0 Bibliography

1. Allan, J. D. (2004). Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. *Annual Review of Ecology, Evolution, and Systematics,35*(1), 257-284. doi:10.1146/annurev.ecolsys.35.120202.110122

2. Ancion, P., Lear, G., Dopheide, A., & Lewis, G. D. (2013). Metal concentrations in stream biofilm and sediments and their potential to explain biofilm microbial community structure. *Environmental Pollution, 173*, 117-124. doi:10.1016/j.envpol.2012.10.012

3. Benoit, G., Wang, E. X., Nieder, W. C., Levandowsky, M., & Breslin, V. T. (1999). Sources and History of Heavy Metal Contamination and Sediment Deposition in Tivoli South Bay, Hudson River, New York. *Estuaries, 22*(2), 167. doi:10.2307/1352974

4. Boëchat, I. G., Krüger, A., Giani, A., Figueredo, C. C., & Gücker, B. (2011). Agricultural land-use affects the nutritional quality of stream microbial communities. *FEMS Microbiology Ecology, 77*(3), 568-576. doi:10.1111/j.1574-6941.2011.01137

5. Dean-Ross, D., & Mills, A. L. (1989). Bacterial Community Structure and Function along a Heavy Metal Gradient. *Applied and environmental microbiology*, *55*(8), 2002-2009

6. Findlay, S. (2010). Stream Microbial Ecology. *Stream Ecosystems in a Changing Environment,* 135-150.doi:10.1016/b978-0-12-405890-3.00003-8

7. Gadd, Geoffrey M., and Alan J. Griffiths. (1978). Microorganisms and Heavy Metal Toxicity. *Microbial Ecology*, vol. 4, no. 4, 1977, pp. 303–317., doi:10.1007/bf02013274

8. Harding, Jon. (2018). Impacts of Metals and Mining on Stream Communities.

9. Hoostal, Matthew & Bidart, M & Bouzat, Juan. (2008). Local adaptation of microbial communities to heavy metal stress in polluted sediments of Lake Erie. FEMS microbiology ecology. 65. 156-68. 10.1111/j.1574-6941.2008.00522

10. Hosen, J. D., Febria, C. M., Crump, B. C., & Palmer, M. A. (2017). Watershed Urbanization Linked to Differences in Stream Bacterial Community Composition. *Frontiers in Microbiology, 8*. doi:10.3389/fmicb.2017.01452

11. Igiri, Bernard E., et al. (2018). Toxicity and Bioremediation of Heavy Metals Contaminated Ecosystem from Tannery Wastewater: A Review. *Journal of Toxicology*,

vol. 2018, pp. 1–16., doi:10.1155/2018/2568038

12. Liess, M., Gerner, N. V., & Kefford, B. J. (2017). Metal toxicity affects predatory stream invertebrates less than other functional feeding groups. *Environmental Pollution, 227*, 505-512. doi:10.1016/j.envpol.2017.05.017

13. Medeiros, J. D., Cantão, M. E., Cesar, D. E., Nicolás, M. F., Diniz, C. G., Silva, V. L., Vasconcelos, A. T., … Coelho, C. M. (2016). Comparative metagenome of a stream impacted by the urbanization phenomenon. *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]*, *47*(4), 835845

14. Qu, Xiaodong & Ren, Ze & Zhang, Haiping & Zhang, Min & Zhang, Yuhang & Liu, Xiaobo & Peng, Wenqi. (2017). Influences of anthropogenic land use on microbial community structure and functional potentials of stream benthic biofilms. Scientific Reports. 7. 10.1038/s41598-017-15624

15. Qu, Xiaodong & Wu, Naicheng & Tang, Tao & Cai, Qinghua & Park, Y-S. (2010). Effects of heavy metals on benthic macroinvertebrate communities in high mountain streams. Annales de Limnologie - International Journal of Limnology. 46. 291-302.10.1051/limn/2010027

16. Rosi, E. J., Bechtold, H. A., Snow, D., Rojas, M., Reisinger, A. J., & Kelly, J. J. (2018). Urban stream microbial communities show resistance to pharmaceutical exposure. *Ecosphere, 9*(1). doi:10.1002/ecs2.2041

17. Sridhar, K & Krauss, Gerd-Joachim & Baerlocher, Felix & Seetharam, Raviraja & Wennrich, R & Baumbach, R. (2001). Decomposition of alder leaves in two heavy metal-polluted streams in central Germany. Aquatic Microbial Ecology - AQUAT MICROB ECOL. 26. 73-80. 10.3354/ame026073

18. Sawkill Watershed and Flood Mitigation Assessment - Final Report. (2018). NEIWPCC and NYSDEC Hudson River Estuary Program

19. C. Urban, Mark & Skelly, David & Burchsted, Denise & Price, William & Lowry, Sarah. (2006). Stream communities across a rural-urban gradient. Diversity and Distributions. 12. 337 - 350. 10.1111/j.1366-9516.2005.00226

20. Utz, Ryan & Hopkins, Kristina & Beesley, Leah & Booth, Derek & Hawley, Robert &

Baker, Matthew & Freeman, Mary & Jones, Krista. (2015). Ecological resistance in urban streams: the role of natural and legacy attributes. Freshwater Science. 000-000. 10.1086/684839

21. Vishnivetskaya, T. A., Mosher, J. J., Palumbo, A. V., Yang, Z. K., Podar, M., Brown, S. D., Brooks, S. C., Gu, B., Southworth, G. R., Drake, M. M., Brandt, C. C., … Elias, D. A. (2010). Mercury and other heavy metals influence bacterial community structure in contaminated Tennessee streams. *Applied and environmental microbiology*, *77*(1), 30211

22. Wilhelm, L., Singer, G. A., Fasching, C., Battin, T. J., & Besemer, K. (2013). Microbial biodiversity in glacier-fed streams. *The ISME Journal, 7*(8), 1651-1660. doi:10.1038/ismej.2013.44

23. Yao, Xie-Feng, et al. (2016). The Effect of Heavy Metal Contamination on the Bacterial Community Structure at Jiaozhou Bay, China. *Brazilian Journal of Microbiology*, vol. 48, no. 1, 2017, pp. 71–78., doi:10.1016/j.bjm.2016.09.007

24. Zeglin, L. H. (2015). Stream microbial diversity in response to environmental changes: Review and synthesis of existing research. *Frontiers in Microbiology, 6*. doi:10.3389/fmicb.2015.00454

25. Zelewski, L.M. & Benoit, G & Armstrong, D.E.. (2001). Mercury dynamics in Tivoli South Bay, a freshwater tidal mudflat wetland in the Hudson River. Biogeochemistry. 52. 93-112. 10.1023/A:1026503215357