MICROBIAL ECOLOGY OF URBAN SEWERS

by

Emily Lou LaMartina

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Freshwater Sciences

at

The University of Wisconsin-Milwaukee

December 2022

ABSTRACT

MICROBIAL ECOLOGY OF URBAN SEWERS

by

Emily Lou LaMartina

The University of Wisconsin-Milwaukee, 2022
Under the Supervision of Dr. Ryan J. Newton

Municipal sewage provides a glimpse into the health and activities of a human society. For more than a century, sewage exploration has helped expose the sources of disease outbreaks and track disease progression over time. Recent advancements in wastewater surveillance born from the COVID-19 pandemic have potential to enhance mitigation efforts against the decades-long global health crisis of microbial antibiotic resistance. However, critical knowledge gaps exist in wastewater surveillance, stemming from a lack of understanding in sewer microbial ecology. Ecology reveals trends in how communities respond and adapt to change, which has far-reaching implications for identifying effective strategies for disease control. However, with little knowledge about sewer microbial communities, including its residents, community dynamics, and functions, no baseline picture of the sewer microbiome exists. The goal of this dissertation was to characterize the sewer microbiome using an ecological approach. The specific aims were to determine if (1) microbial communities in urban wastewater exhibit seasonal patterns in assembly; (2) if seasonal community assembly is driven by environmental bacteria responding to changes in water temperature; and (3) if temperature-driven communities modulate the composition and abundance of antibiotic resistance genes in wastewater. Results show that microbes in sewers have seasonal

community dynamics akin to other natural environments, and they have adapted to this stressful environment by acquiring and maintaining mechanisms of antibiotic resistance. Using only well-established methods in DNA sequencing and analyzing a wastewater dataset covering expansive temporal and spatial scales, this dissertation builds the foundation of a baseline sewer microbiome in the United States. All data collected and analyses used were made publicly available to aid standardizing methods in global strategy plans. Together, standardizing methods and sharing data related to the sewer microbiome will improve predictive models, guide interventions, and make other public health breakthroughs in wastewater surveillance.

For Mila

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Sewage: a nexus of the urbanized and natural worlds

In 1854, a deadly illness swept across London, England, that claimed 500 lives in the first 10 days[1]. Patients suffered from a distinctive "rice water" diarrhea that led to severe dehydration and oftentimes death. Breathing air from the squalid river Thames was blamed as causing the disease, om a now-defunct theory known as *miasma*. A local anesthesiologist, John Snow, rejected *miasma* and sought to identify its true source. Snow placed dots on a map of London at residences of afflicted persons. The density of dots exposed disease hotspots that clustered around water pumps. With this discovery, Snow argued that diseased waste was contaminating the London water supply and spreading the epidemic[1].

The culprit to what is now known as the Broad Street cholera outbreak of 1854 was a bacterium that to this day plagues communities with limited access to clean water – *Vibrio cholerae*. Snow used the distribution and frequency of disease occurrences to track its source. This was a substantial feat, considering Louis Pasteur did not discover that some diseases are caused by microorganisms for another ten years[2]. Wastewater ultimately became a vital tool for epidemiology, one that takes snapshots of the guts of a human society. Today, it can help monitor public calamities such as the frequency of disease[3], obesity rates[4], and illicit drug use[5]. The utility of wastewater monitoring came especially to light during the COVID-19 pandemic[6,7], once its potential for tracking the SARS-CoV-2 virus was proposed[8,9], implemented[10], and optimized[11,12]

to such a degree that population-wide occurrences could be predicted[13,14], leaving lasting improvements on water distribution[15,16] and treatment[17,18]. With a widespread monitoring effort established, it has been shown that disease outbreaks can be mitigated before disaster strikes. This was recently demonstrated in New York City, where wastewater monitoring revealed the presence of the polio virus, the first time it had been detected in thirty years[19].

The concepts of disease caused by microorganisms may be relatively new to human history, but sewers, storm drains, and other water sanitation systems have existed for millennia. Around 600 BC, the Romans engineered *Cloaca Maxima*, a sewer that directs waste from the city to the River Tiber[20]. Water sanitation in the United States is comparatively young, but its progress has been exponential. Chicago and Brooklyn were the first to build sewerage systems, starting in the 1850s[21]. In 150 years, sewers in the US had expanded to more than 1.3 million miles[22]. Each day, 34 billion gallons of sewage is collected and diverted to wastewater treatment plants (WWTPs)[23], but the US infrastructure, now inadequate and corroded, leaches 7% of untreated sewage into the environment each year[24]. Contaminated water is linked to 7.2 million illnesses annually[25]. In 1993, in Milwaukee, WI, *Cryptosporidium* in drinking water caused more than 400,000 gastrointestinal illnesses, 4,000 hospitalizations, and 100 deaths[26].

Throughout most of history, wastewater treatment could be summarized with the adages, “the solution to pollution is dilution” and “out of sight, out of mind.” As societies developed, humans dumped their waste in local waterways[20] and buried it in cesspools[27]. During the agricultural revolution, sewage was used to water and fertilize farmland[28]. Urbanization soon followed, and

with people densely packing into cities, this type of sanitation was no longer sufficient (a la, the Broad Street cholera outbreak). Scotland built the first municipal sewage treatment plant in the early 1800s, utilizing a sand filtration system[29]. Fifty years later, in Kent, UK, chlorine bleach was added to the water main to control a typhoid outbreak[30], the first use of chlorine to sanitize a public water.

Microorganisms majorly contribute to wastewater treatment in two distinct compartments, activated sludge and anaerobic digesters[31]. In *activated sludge*, wastewater is aerated to allow microbes to oxidize organic materials, breaking them down and forming flocs. Flocs settle in tanks to facilitate solids filtration. *Anaerobic digestion* is a specialized microbial process that breaks down organic matter for energy in the absence of oxygen. Anaerobic digesters in WWTPs cultivate these special communities to convert waste, such as food, oil, and manure, into biogases, mainly methane and carbon dioxide[32]. Advanced facilities can conserve energy and reduce greenhouse gas emissions by using sequestered biogases as a power supply[33,34]. Milwaukee, WI, was actually the first city to use anaerobic digestion in a full-scale WWTP and resell its sludge as a garden fertilizer, *Milorganite* (*Milwaukee organic nitrogen*)[35]. Together, filtration (primary treatment), activated sludge (secondary treatment), and chlorination (tertiary treatment) make up the foundation of modern wastewater treatment plants, which boast a pathogen removal efficiency of 99%[36].

Microbial ecology and its paradigms

Anaerobic digestion is one example of many that showcase how microorganisms influence life on earth. At a global scale, microbes contribute some part to all biogeochemical cycles. Microbes biologically convert atmospheric nitrogen (N_2) to ammonia (NH_3) and strip phosphates (PO_4^{3-}) during decomposition, offering essential biosynthesis building blocks for all other organisms. Bacteria in the ocean produce 20% of atmospheric oxygen (O_2)[37], and their role in the sulfur cycle even helps the formation of clouds[38]. The study of these microbial processes, along with the interactions between the microbes and their environment, is known as microbial ecology.

Microbial ecology is an important topic in both human and environmental health. Microbes are major forces in climate change because they can produce and consume greenhouse gases. In medicine, critical connections have been made between human health and the microbes that inhabit the body, or the microbiome. For example, in a fecal transplant, the gut microbiome from a healthy person colonizes the gut of a patient whose microbiome has been wiped out by a *Clostridium difficile* infection, effectively curing the disease[39]. Chronic conditions such as obesity[40] and depression[41] can also be linked to the structure of a microbiome. A person carries more bacterial cells than human cells[42], so the potential for further medical advancements coming from microbiome research seems limitless.

Microbial ecology relies on empirical evidence. Observations generate questions, questions are framed as hypotheses, then experiments are designed to test the hypotheses. There are a multitude of variables that can cause any given effect, so empirical evidence is used to identify thresholds,

metrics, and indicators that best reflect the function of interest[43]. These variables are often arbitrary but nonetheless necessary to emphasize some effects while ignoring others. For example, *Escherichia coli* is a commonly used indicator of risk related to beach closures. *E. coli* is abundant in human feces, therefore its presence in beach water can indicate sewage pollution. There are thousands of different fecal bacteria that could have been chosen, but *E. coli* was the most applicable because it is easy, fast, and inexpensive to quantify in a lab. Over time, data collected on *E. coli* levels, weather patterns, hospital visits, and other metrics have accumulated. Mounting empirical evidence and advancements in technology are improving beach-closure risk assessments, thus more accurate indicators of human fecal pollution have been proposed[44–46].

DNA sequencing technology is integral to microbiome research. It translates the arrangement of constituent DNA molecules (i.e., the deoxynucleotides deoxyadenosine, deoxythymidine, deoxycytosine, and deoxyguanosine) to legible symbols (i.e., letters A, T, C G). First-generation sequencing in the 1970s was laborious, slow, and could only output the composition, not order, of one-thousand nucleotides[47]. Despite the challenges, significant medical breakthroughs were made in just a few years, including the genetic origin of Huntington's disease[48]. Second-generation sequencing dramatically improved efficiency by recording nucleotides during DNA synthesis, omitting the need for arduous, follow-up visualization techniques. Currently, so-called next-generation sequencing has been so systematically optimized, the rate it generates data greatly exceeds what was considered technically possible[47,49,50]. Of course, with increased accessibility comes decreased cost, and a prime example of this was demonstrated by the Human Genome Project (genome.gov/human-genome-project). In 2003, the complete human genome with 3 billion

nucleotides was sequenced, a triumph that took 13 years and cost 3 billion dollars. Ten years later, a human genome could be sequenced in 2 weeks for only one-thousand dollars[51].

In ecology, genomic differences are used to study phylogeny, or relatedness, of genes and organisms based on their inferred evolutionary distance from a common ancestor. When dealing with a group of closely related organisms, comparing whole genomes may be impractical, given that most DNA sequences will be shared (human and chimpanzee genomes, for example, are 99% identical[52]). Marker genes are ones shared by all organisms in a group of interest and are a more practical tool for assessing phylogeny at a community level. Phylogenetic distance between organisms increases as sequences in hypervariable regions in marker genes become less similar. Study of the microbial marker gene proposed in 1977 by Woese and Fox[53] has greatly expanded knowledge of the tree of life[54]. The 16S rRNA gene encodes the small subunit of ribosomes in bacteria and archaea, and hypervariability in non-conserved loop regions create nine distinct regions, V1 through V9, that are useful for community analysis. Taxonomic cutoffs for microbial groups are harder to discern than for organisms that sexually reproduce, so the limitations of 16S rRNA marker genes are debated[55–57]. Nevertheless, recent opinions support that sub-regions can effectively differentiate species-level differences ($\geq$ 97% similarity) and sequencing the entire 16S rRNA gene can allow strain-level comparisons (> 99%)[58]. For relatively low cost and minimal computational demand, 16S rRNA gene sequences infer taxonomy at resolutions fine enough to reveal complex and influential ecosystem dynamics[59–63].

The antibiotic resistance crisis: an ecological perspective

Since the discovery of penicillin almost 100 years ago, antibiotics have added 23 years to the average human lifespan[64]. In the beginning, new antibiotics were regularly being discovered and applied without prudence. However, new discoveries peaked by the 1950s, and the current arsenal of antibiotics has provided minimal defense against the escalating cases of resistant superbugs. By 2001, the World Health Organization declared antibiotic resistance a top global health threat[65,66] and promising therapies were kept closely guarded and controlled. However, the efficacy of novel treatments is still relentlessly brief. In the recent case of daptomycin, an *Enterococcus faecium* infection was unperturbed by treatment only two years after the drug was introduced[67].

The term antibiotic is actually a misnomer[68]. There are few ecological examples of biomolecules having antibiosis functions[69] and naturally-occurring antibiotics in soil and marine habitats exist at such a low levels, growth inhibition is rarely achieved[70,71]. Sub-lethal levels have been shown to trigger multiple metabolic pathways in bacteria, so it is possible that antibiotics have alternative ecological roles, such as cell-cell signaling[72–74]. For example, in quorum sensing, pathogens pick up signals from neighboring cells that at a certain density will initiate a host infection. Signal molecules interact with metabolic machinery in a cell, which regulates the expression of virulence genes. A study by Goh et al.[75] produced similar metabolic responses from antibiotics. In the study, *Salmonella typhimurium* and *E. coli* were exposed to sub-lethal doses of rifampicin and erythromycin. Exposure repressed virulence gene expression in *S. typhimurium* and *E. coli* that were resistant to the antibiotics, and conversely, virulence gene expression was activated in those susceptible to the antibiotics. Therapeutically, rifampicin and erythromycin treat infections by

directly disrupting metabolic machinery (RNA polymerase and ribosomes, respectively). This suggests that, at high doses, these antibiotics inhibit growth by interfering with cell metabolism, but at environmentally relevant doses, they regulate metabolism by interacting with cell machinery. Exploring alternative ecological roles of these compounds was often understandably overlooked at the onset of antibiotic therapies, because of the immediate value of these 'miracle drugs' that can eliminate deadly infections without harming the human host.

An important metabolic pathway that responds to external signals is the microbial stress response. The stress response encompasses a cascade of metabolic changes that help populations adapt to changing environments. Normal functions are swapped out and replaced by ones that enable the genome to diversify, which in turn can facilitate the development and dissemination of antibiotic resistance genes (ARGs). Stress responses with potential to propagate ARGs include (1) arresting error-correcting DNA polymerases, increasing the likelihood of favorable mutations on antibiotic target sites[76]; (2) accelerating horizontal gene transfer (HGT) of plasmids carrying ARGs[77–79], even between distantly related taxa[80]; (3) reshuffling gene cassettes on mobile genetic elements (MGEs), bringing ARGs closer to the promoter and increasing their expression[81]; (4) triggering biofilm formation, where the dense packing of cells allows HGT to occur more readily, and where an extracellular matrix creates a protective barrier against antibiotics[82].

Given these features of the microbial stress response, it is possible for novel ARGs to develop in stress-inducing environments, including those contaminated with antibiotics at sub-lethal levels[83,84]. However, certain obstacles must be overcome (in no specific order) for environment-

derived ARG to take hold in a human pathogen. First, in order for an ARG to become fixed in a community, it faces a founder effect, where its function must not only be new, but required[85]. Second, the benefit of the ARG must outweigh the cost of maintaining and expressing it. Third, the ARG must become mobilized on a plasmid or MGE capable of transfer from the environmental bacterium to a human pathogen. Finally, the pathogen must have the ability to translate the ARG to a functioning protein[86].

Sewage offers routes that circumvent obstacles between environment-derived ARGs and pathogens. Sewers collect antibiotic residues in their active form in the waste of humans and animals receiving treatment. These residues can trigger stress responses that allow mutations, which can lower the cost of expressing the ARG and promote its dissemination in the community. Human pathogens that use sewers as a main habitat carry the greatest risk for propagating ARGs, since they can mediate ARG transfer between environmental bacteria and other pathogens[86]. Members of the genera *Arcobacter*[87] and *Aeromonas*[88] are emerging pathogens worth noting that are ubiquitous in urban waters and have exhibited multidrug resistance[87,88].

Despite their potential importance as vectors for ARGs[89–92], influencers of wastewater treatment[93], polluters of recreational waters[94], and corroders of sewer pipes[95–99], little is known about microbial communities that use sewers as a main habitat. Only recently were sewer microbial communities characterized[100] and proposed as ecosystems ten years later[101]. Within WWTPs, microbial communities have been extensively studied[102–106] but fewer have focused on pre-treatment, resident sewer microbes. Previous work identified major taxa that consistently comprise these

communities[4,100], but without further investigating sewage collected across broad space and time intervals, whole community dynamics remain unclear.

It is not yet known if sewer microbial communities have a stable structure, fluctuate stochastically, or change in response to environmental stimuli. This lack of understanding hinders our ability to determine the fate of microbes in sewage pollution, identify those most responsible for spreading antibiotic resistance, or understand how the environment influences microbial functional capacity, particularly related to risk to human and animal health. Moreover, insight gained from wastewater monitoring is limited without an established "baseline" community. It is impossible to confidently suggest that one variable is the major driver of an observation without first understanding the fundamental dynamics of the system.

Aims, research strategy, and significance of this dissertation

This dissertation defines sewer microbial ecosystems using ecology fundamentals: whole community dynamics, community functions, and major community members. The goal was to discern a baseline picture of the sewer microbiome. To achieve this goal, wastewater samples representing broad temporal and spatial scales were analyzed with microbial marker gene and metagenomic DNA sequencing. In the end, major taxonomic groups were characterized; predictable community changes over time and space were revealed; and consistencies and deviations of ARGs in relation to predictable community changes were identified. The specific aims of this research were to:

1. **Determine if seasonal community assembly is driven by environmental taxa that respond to changes in wastewater temperature.**
    a. Show seasonal clustering of samples in a 16S rRNA gene sequence dataset from a multi-year time series of pre-treatment wastewater.
    b. Parse human-associated sequences to separate the resident-sewer bacterial community.
    c. Identify bacterial groups that are indicators of each seasonal steady state.
    d. Test hypothesis that indicator groups are environmental/resident-sewer taxa.
    e. Test hypothesis that relative abundances of indicators are predictable each year, whereas human-associated groups are not, in time series analysis.
    f. Design quantitative assays for droplet digital PCR (ddPCR) to analyze fluctuations of indicator groups observed in sequence data.
    g. Test the hypothesis that temperature also influences community assembly in a dataset covering significant geographic distances and multiple time points.

2. **Determine if temperature-driven microbial communities modulate the composition and abundance of antibiotic resistance genes in wastewater.**
    a. Mine ARGs from pre-treatment wastewater metagenomic datasets from multiple locations and time periods.
    b. Test hypothesis that locations harbor unique pools of ARGs.
    c. Test hypothesis that ARGs in those pools fluctuate over time in patterns that reflect seasonal changes in the host community.

d. Clarify that identified wastewater ARGs are clinically relevant.
e. Quantify ARGs with ddPCR to support abundance changes appear coupled to the warm-dominating microbial community.

The studies addressing these aims (**Aim 1, Chapter II and Aim 2, Chapter III**) are foundational to establishing a baseline understanding of the sewer microbiome. In **Aim 1, Chapter II**, temporal variability was captured by sequencing 16S rRNA genes from wastewater samples collected monthly from two WWTPs in Milwaukee, WI. Spatial variability of wastewater microbial communities was captured by sequencing 16S rRNA genes from samples collected from WWTPs across the USA, as well as manholes from several neighborhoods in Milwaukee, WI. The resident sewer community was estimated by determining 16S rRNA gene sequences that were not found in the Human Microbiome Project database. To support compositional changes observed with DNA sequence analysis, quantitative PCR assays were designed to target bacterial groups that exhibited seasonal abundance patterns. A follow-up resource announcement with full-length 16S rRNA gene sequences from select samples is described in **Appendix A**. **Aim 2, Chapter III** used metagenomic sequencing to mine all possible ARGs from wastewater samples from several locations and time points. Again, quantitative PCR supplemented DNA sequence analysis by showing seasonal abundance patterns of ARGs.

This dissertation also sets up future aims to compare genomes of a major resident sewer group. **Aim III, Appendix B** describes progress towards a pangenome analysis of *Flavobacteria*:

1. **Determine if resident-sewer *Flavobacteria* are genetically distinct from relatives in other freshwater environments.**
    a. Isolate and cultivate bacteria in non-targeted freshwater medium (R2A) from pre-treatment wastewater during the fall and spring community steady states.
    b. Choose isolate candidates based on colony morphology with known *Flavobacteria* characteristics (e.g., yellow-orange color, gliding on surfaces).
    c. PCR-screen candidate isolates using custom sewer-specific *Flavobacterium* primers.
    d. Sequence full-length 16S rRNA genes from isolates identified as *Flavobacterium*.
    e. Select isolates with most distinct 16S rRNA gene sequences to undergo full-genome sequencing.
    f. Compile other freshwater *Flavobacteria* genomes from public genome repositories.
    g. Conduct pangenome analysis of *Flavobacteria* genomes.
    h. Highlight genomic similarities and differences of sewer *Flavobacteria* to its relatives.

Appendix C lists collaborations completed during the years of dissertation research. All of these projects honed skills necessary for this dissertation, particularly in DNA sequence analysis.

CHAPTER II

URBAN WASTEWATER BACTERIAL COMMUNITIES ASSEMBLE INTO SEASONAL STEADY STATES

Published in:

Microbiome

DOI: https://doi.org/10.1186/s40168-021-01038-5

Emily Lou LaMartina[1], Aurash A. Mohaimani[2], Ryan J. Newton[1]*

[1] School of Freshwater Sciences, University of Wisconsin-Milwaukee, 600 E Greenfield Ave., Milwaukee, WI 53204

[2] Current address: Microarray Bioinformatics, Genetic Sciences Division, ThermoFisher Scientific

* Corresponding Author: newtonr@uwm.edu

(*Received 11 April 2020, Accepted 17 February 2021*)

Abstract

Microorganisms in urban sanitary sewers exhibit community properties that suggest sewers are a novel ecosystem. Sewer microorganisms present both an opportunity as a control point for wastewater treatment and a risk to human health. If treatment processes are to be improved and health risks quantified, then it is necessary to understand microbial distributions and dynamics within this community. Here, we use 16S rRNA gene sequencing to characterize raw influent wastewater bacterial communities in a 5-year time series from two wastewater treatment plants in Milwaukee, WI; influent wastewater from 77 treatment plants across the USA; and wastewater in 12 Milwaukee residential sewers. In Milwaukee, we find that in transit from residences to treatment plants, the human bacterial component of wastewater decreases in proportion and exhibits stochastic temporal variation. In contrast, the resident sewer community increases in abundance during transit and cycles seasonally according to changes in wastewater temperature. The result is a bacterial community that assembles into two distinct community states each year according to the extremes in wastewater temperature. Wastewater bacterial communities from other northern US cities follow temporal trends that mirror those in Milwaukee, but southern US cities have distinct community compositions and differ in their seasonal patterns. Our findings provide evidence that environmental conditions associated with seasonal change and climatic differences related to geography predictably structure the bacterial communities residing in below-ground sewer pipes.

Introduction

Urban sewers collect wastewater from a variety of sources, including stormwater, industrial waste, and residential sewage. Sewer pipes transport wastewater to wastewater treatment plants (WWTPs), where nutrients and microorganisms are removed and select microorganisms are cultivated to aid treatment processes[107]. Imbalanced WWTP microbial communities can disrupt treatment and create challenging and costly problems. For instance, WWTPs typically settle activated sludge to separate it from treated wastewater, but overgrowth of filamentous bacteria causes poor settling, which deteriorates effluent quality and may require significant process alterations to remedy[108]. The goal of wastewater treatment is to foster beneficial microbial communities and remove problematic ones, and WWTP influent can be a source of each[3-5].

Sewers serve as more than conveyance for wastewater. The consistency in sanitary sewer microbial community composition suggests that sewers represent a recently formed ecosystem[112]. Some resident sewer microbes induce pipe corrosion[95,99], display pathogenic lifestyles[113], or propagate antibiotic resistance genes[114,115], including those that survive treatment and persist in receiving waters[12-15]. Aging and inadequate infrastructure also introduces sewer bacteria to the environment by leaching wastewater through corroded pipes[16-18] or through deliberate release during sewer overflows[122,123]. Sewage discharge regularly impairs recreational waters, causes coastal beach closures, and poses a significant risk to human health[124]. Despite the potential importance of resident sewer bacteria, there is not a thorough understanding of whether most of these microorganisms exhibit predictable abundance patterns through time or among sewer

systems, partition to various substrates in wastewater, or survive for prolonged periods in natural aquatic systems after discharge.

Many aquatic ecosystems undergo seasonal changes that drive biological change, which in turn creates repeating and predictable microbial community structures and ecosystem services[22-25]. As sewers are a primarily aquatic environment, it is possible the resident microbial communities also exhibit temporal community assembly patterns. Initial studies suggest this may be the case. Guo et al.[129] revealed diurnal trends in WWTP influent microbial communities that were driven by change in flow rate between day and night, where low flow resulted in less sloughing of pipe bacteria and thus a change in composition. Although this study provided evidence of repeatable microbial dynamics, these dynamics were driven by short-term physical factors. To the best of our knowledge, no study has analyzed whether pre-treatment wastewater microbial communities are also impacted by longer-term changes (months or years) to their environment. Uncovering patterns of assembly by sewer microbial communities will aid in designing models to predict wastewater composition, enable targeted treatments for microorganisms of interest, and identify whether temporal community variation relates to altered human and/or environmental health risks from untreated discharge.

To address this knowledge gap, we used 16S rRNA gene sequencing to analyze bacterial communities in three wastewater datasets: (1) a 5-year time series of WWTP influent sampled once per month from two facilities in Milwaukee, WI, USA; (2) WWTP influent from 77 facilities in the USA sampled during three seasons in a single year; and (3) wastewater from 12 sewers in

four distinct residential Milwaukee neighborhoods. To assess the mixing of microbiomes, or the "community within a community," we identified the human-associated bacterial assemblage in wastewater and analyzed it independently from the rest of the bacterial community. We hypothesize that (1) most sewer pipe bacteria are not from human waste and persist year-round; (2) wastewater resident bacterial communities follow predictable, seasonal patterns in assembly; and (3) temporal community assembly trends in Milwaukee will be similar to wastewater from other northern US cities.

Material and Methods

Sample collection and DNA extraction

Milwaukee time series

We collected 24-h flow-proportional composite samples of WWTP influent once a month for 5 years from Jones Island (JI) and South Shore (SS) water reclamation facilities in Milwaukee, WI, USA (Table 1). At JI, 100 mL aliquots from continuous water sampling at three sample points were combined into a final composite sample. Each sampling point has variable sampling frequencies depending on flow at that location. Under low flow (range = 10 million gallons per day (MGD) to 120 MGD depending on sample point), the volumes that trigger an aliquot collection are (1) 0.2 MG, (2) 0.5 MG, and (3) 0.6 MG, respectively, while under high flow (range => 60 MGD to > 120 MGD) the volumes that trigger an aliquot collection are (1) 0.8 MG, (2) 1.0 MG, and (3) 1.4 MG. At SS, a single composite sample was collected. Under low flow conditions (< 100 MGD) a 100-ml aliquot is collected at 0.7-1.9 MG, while under high flow conditions (≥ 100

MGD) an aliquot is collected at 1.9-4.0 MG. JI influent samples spanned each month from January 2013 to February 2018, except 2 months (November 2014 and March 2015; n = 60). SS influent samples spanned October 2014 to December 2017, except 5 months (November 2014, March 2015, May 2015, November 2015, and June 2017; n = 34). After collection, we filtered 10-ml onto 0.22-μm mixed cellulose ester filters (47-mm diameter, Millipore Sigma) and stored at -80 °C for up to 5 years before extracting DNA. The Milwaukee Metropolitan Sewerage District measured environmental parameters in each sample.

Milwaukee residential sewers

We collected 5-h time-paced composite (0400-0900 h, with 50 mL aliquots taken every 15 min) samples from three sewers in each of four neighborhoods, Elm Grove, South Milwaukee, North Milwaukee, and New Berlin in the Milwaukee sewerage district on the 15th and 17th of December 2015 (n = 24; Fig. 1 and Table 2). Each residential sewer sample represented a 200-600 household drainage area. From these samples, we filtered 25-ml onto 0.22-μm mixed cellulose ester filters (Sigma Millipore) and stored them at -80 °C for up to 3 months before extracting DNA.

Across USA

As described previously in Newton et al. 2015[4], sewage influent samples (n = 204) were collected from 77 wastewater treatment plants (WWTPs) in 72 US cities around August 2012, January 2013, and May 2013 (Fig. 2 and Table 3). Wastewater samples included a variety of collection setups, ranging from single time-point grab samples to 24-h flow weighted composites. All samples were collected, stored in a refrigerator on site for < 24 h and shipped overnight to our lab for immediate

filtering onto 0.22-μm mixed cellulose ester filters. For specific sample collection details, see Newton et al. 2015[4].

For each sample set, we crushed frozen filters in their storage tubes using a sterile metal spatula, added a bead-beating matrix and buffers from the FastDNA Spin Kit for Soil (MP Bio), and bead beat for 1 min. We then extracted DNA following the FastDNA Spin Kit for Soil protocol.

PCR and amplicon sequencing

We amplified the V4-V5 region of bacterial 16S rRNA genes in wastewater samples using primers 518F and 926R[130]. The following setup was used: 12.5-μl 2x KAPA HiFi HotStart ReadyMix PCR (Roche), 1.5-μl of each 5-μM forward and reverse primer working solutions, 7.5-μl sterile water, and 2-μl 100x-diluted DNA template. PCR was run on a vapo-protect Mastercycler pro S (Eppendorf) under the following conditions: 95 °C for 5 min; 22 cycles of 98 °C for 20 s, 55 °C for 15 s, 72 °C for 1 min; 72 °C for 1 min; 4 °C hold. We included one negative control (PCR blank) and one mock community (#HM-782D, BEI). Triplicate PCRs were pooled and cleaned with Agencourt AMPure XP beads (Beckman Coulter), following the manufacturer's protocol. Sample libraries were prepared according to the Illumina MiSeq protocol in the Nextera XT Index kit (Illumina). Indexed PCR amplicons were cleaned with AMPure beads and normalized with the SequalPrep Kit (ThermoFisher). Sequencing was carried out on an Illumina MiSeq with 2 x 250 chemistry at the UW-Milwaukee Great Lakes Genomics Center (UWM GLGC; Research Resource Identifier: SCR_017838; greatlakesgenomics.uwm.edu) for the Milwaukee time series samples and the Marine Biological Laboratory for the residential sewers and USA samples.

Sequence processing

Forward and reverse reads were quality-filtered using FastQC[131] and primers were trimmed with Cutadapt[132]. We processed the three wastewater datasets simultaneously with the R package DADA2[133], with the following specifications: during filtering, reads were truncated at 230 bp, and reads with quality scores lower than 10 were removed; after merging, sequences were removed that did not have lengths within 5% (355 to 393) of the median sequence length (374 bp). Taxonomy was assigned to resulting amplicon sequence variations (ASVs) using SILVA v.132[134]. ASVs that were not classified as bacteria or were classified as mitochondria or chloroplasts were removed. Contaminant ASVs from the mock community and negative control were identified with the R package decontam[135] and subsequently removed.

Primer design

We designed primers to target unique 16S rRNA V4-V5 gene regions belonging to one *Cloacibacterium* and one *Flavobacterium* ASV (Table 4). Non-target ASVs of the same genus were included as negative controls for primer design and PCR amplification. MEGA7[136] was used to align target and non-target sequences and identify the most variable regions for primer design. We used Primer3[137] to design primer sequences and calculate annealing conditions. Target specificity was checked against RDP Probe Match[138].

Gene quantitation

Quantitative PCR (qPCR)

Target ASVs were quantified in all 60 samples of the JI time series using droplet digital PCR (ddPCR). Reactions were set up as follows: 11-µl EvaGreen Supermix (Bio-Rad), 1.3-µl 5-µM forward and reverse primers, 6.4-µl sterile water, and 2-µl 100x-diluted DNA template. PCR was run on a vapo-protect Mastercycler pro S under the following conditions: 95 °C for 5 min; 40 cycles of 95 °C for 30 s, 58-60°C for 1 min; 4°C for 5 min; 90°C for 5 min; 4°C hold. The human *Bacteroides* marker[139], a human fecal marker in the genus *Bacteroides*, was quantified in the first 48 samples of the JI time series using qPCR following methods described previously[45].

Droplet digital PCR (ddPCR)

Target ASVs (flavo11, flavo42, cloaci08, cloaci32) were quantified in all 60 samples of the JI time series using droplet digital PCR (ddPCR). Reactions were set up as follows: 11-µl EvaGreen Supermix (Bio-Rad), 1.3-µl 5-µM for- ward and reverse primers, 6.4-µl sterile water, and 2-µl 100x-diluted DNA template. PCR was run on a vapo- protect Mastercycler pro S under the following conditions: 95°C for 5 min; 40 cycles of 95°C for 30 s, 58-60°C for 1 min; 4°C for 5 min; 90°C for 5 min; 4°C hold. Gene blocks of V4-V5 sequences (IDT) were used as positive controls (Table 5).

Partition reads from the human microbiome

We pulled Human Microbiome Project (HMP) studies 16S-PP1 and 16S-PP2 from the HMP resource page (https://hmpdacc.org/hmp/). HMP sequence IDs were uniquely de-replicated by their URL address and concatenated into a single FASTA. A tool was created to reduce unique sequences that occurred at least 10 times with at least one subject and sample ID available for

each, with at least 1 subject and sample ID available for each uniquely filtered sequence. A parallelized, exact-identity sequence aligner was implemented and employed to align the sewage ASVs against the curated HMP reference database. We identified 491 human-associated ASVs within 35,332 total wastewater ASVs. For this study, the remaining ASVs were considered sequences from resident sewer microorganisms. Human-associated ASVs were binned by source body site (Table 6).

Due to sequencing errors and potential microorganism transfer among source environments, we established a threshold to identify and partition low-abundance, uncommon human-associated reads that were common in sewer samples. Among ASVs that were shared between WWTP influent and the HMP, if the minimum relative abundance across samples (5th-percentile) of a wastewater ASV exceeded the maximum relative abundance across samples (95th-percentile) of that ASV in the human microbiome, it was reclassified as a sewer-associated sequence. If the minimum abundance (5th-percentile) of a wastewater ASV was less than the maximum abundance (95th-percentile) of that ASV in the human microbiome, it was considered a human-associated sequence. After the filtering procedure, we moved 33 ASVs from a human-associated to a sewer-associated classification. In the final dataset, 458 ASVs in the wastewater samples were classified as human-associated.

Statistics and graphics

The Shannon diversity index, a measure of alpha diversity, Bray-Curtis dissimilarity, a measure of beta diversity, and ordinations were calculated using the R package vegan[140]. We also used Mann-

Whitney U tests, hierarchical clustering, autocorrelation function, linear regression, Shapiro-Wilk tests, and ANOVA to examine statistical relationships in the data, and these were performed using the R stats[141] package. We identified indicator ASVs with the R package indicspecies[142]. To reduce dataset complexity and examine the predominant bacteria, only ASVs with a maximum relative abundance of 1% or greater were considered in the indicator analyses. Principal coordinate analyses (PCoA) were conducted with the R package ape[143]. All figures were made in R with ggplot2[144].

More specifically, we performed the following analyses to visualize and/or test statistically for differences in the community composition and abundance (qPCR/ddPCR) datasets: (1) a non-paired Mann-Whitney U test to compare Shannon diversity values between the US city and Milwaukee WWTP influent time-series datasets and the Milwaukee neighborhood and WWTP influent time-series datasets, 2) a non-paired Mann-Whitney U test to compare Bray-Curtis dissimilarity values between the US city and Milwaukee WWTP influent time-series datasets and the Milwaukee neighborhood and WWTP influent time-series datasets, (3) a Principle Coordinate Analysis (PCoA) of the Milwaukee time-series dataset to examine temporal patterns in community composition, (4) an indicator analysis (indicspecies[142]) to identify ASV relative abundance patterns that are indicative of groups of months, here set at exactly three consecutive month groupings in the Milwaukee time-series dataset, (5) a PERMANOVA test to identify if the month-based seasonal groupings of community composition are different statistically in the Milwaukee time-series dataset, (6) a correlation of environmental and sample metadata to the Bray-Curtis dissimilarity of bacterial community composition (envfit in R package vegan[140]) in the Milwaukee

time-series dataset, (7) an examination of seasonality in abundance patterns of individual ASVs using hierarchical clustering of z-score normalized ASV relative abundances, where each ASV was relativized to its relative abundance values across the Milwaukee time-series samples, (8) the autocorrelation function (ACF) with 60 1-month time lags to test for data self-similarity with a defined time-lag; i.e., a test of significant seasonal patterns in the relative-abundance of particular ASVs, (9) Spearman rank correlations to test for relationships between ASV relative abundance data and the quantitative PCR data for select ASVs, and (10) a Mann-Whitney U test for seasonal differences in the US city community composition data (e.g., cold period northern city vs. warm period southern city). For more detailed information on the specific functions used, see Table 7 or visit https://github.com/NewtonLabUWM/Sewage_TimeSeries.

Results

Wastewater bacterial community diversity scales with time and space

The Shannon diversity index was similar between samples collected in the Milwaukee time series and US city WWTP influent datasets (Mann-Whitney U, $p = 0.33$; Fig. 3a), which indicates there is a relatively consistent number/evenness of bacterial taxa that co-inhabit these municipal sewer systems. In contrast, Shannon diversity was greater in the Milwaukee neighborhood wastewater samples than in the Milwaukee WWTP influent samples (Mann-Whitney U, $p = 4.9 \bullet 10^{-11}$; Fig. 3a).

Contrasting the alpha diversity measure, the bacterial community composition was not similar across the WWTP influent datasets. Bray-Curtis dissimilarity increased as the sample set included

more WWTPs or time points covering a greater proportion of a year in a single treatment plant (Fig. 3b). The range of Bray-Curtis dissimilarity values was similar between residential wastewater samples and the Milwaukee WWTP influent samples (Mann-Whitney U, $p = 0.12$) but was greater in the US city dataset (Mann-Whitney U, $p < 2.2 \cdot 10^{-16}$). This result indicates that differences in environmental conditions among sewers have a larger influence on community composition than localized, within-system environmental differences.

Resident sewer bacterial communities are distinct from the human microbiome

Bacteria associated with human stool became a lesser part of the overall community as wastewater traveled from neighborhood sewers to the WWTP (Fig. 4a). For example, *Bacteroides* was on average the most abundant genus in human stool (53% of community). It decreased to 11% in Milwaukee residential sewer communities and 3.4% in Milwaukee WWTP influent. *Acinetobacter* was the most abundant genus in WWTP influent communities in Milwaukee (11% of community) and across the US (8.8%). *Acinetobacter* was not as dominant in residential wastewater (5.3%) and was virtually absent ($4.5 \cdot 10^{-3}$ %) in human stool. Other abundant stool-associated genera, including *Alistipes*, *Faecalibacterium*, and *Parabacteroides*, also decreased in their contribution to the overall community as they moved from the human host, into the sanitary sewer system, and to the WWTP. Their dominance was replaced by genera not common to the human microbiome, such as *Arcobacter*, *Trichococcus*, and *Flavobacterium* (Fig. 4b).

The majority of wastewater bacteria were not associated with the human microbiome (Fig. 4c). In residential sewer communities, 35.9 ± 7.5% of reads belonged to ASVs attributed to the human

microbiome, but in the 5-year time series of two Milwaukee WWTPs, the proportion dropped to 11.0 ± 2.8%. Similarly, across the USA, only 12.4 ± 5.7% of reads were human-associated. Of the human microbiome sources, stool was the greatest contributor of ASVs to WWTP influent (9.0 ± 4.7%; Fig. 4d). Overall, we find that the majority of reads in wastewater were assigned to ASVs that were not associated with the human microbiome (88.0 ± 5.0%), and we considered them to be sewer-associated for subsequent analyses.

Wastewater bacterial communities assemble into seasonal steady states

Milwaukee WWTP influent bacterial communities repeatedly assembled into two community states each year (Fig. 5a), and the pattern was consistent for both of Milwaukee's WWTP facilities (Fig. 5b). In a PCoA, all samples from January through May had Axis 1 scores less than 0, while samples from August through November had Axis 1 scores greater than 0. Samples from June, July, and December had both positive and negative Axis 1 scores. Typically, samples from April to May and September to October harbored the most distinct community compositions (Fig. 5a).

We conducted an indicator analysis to identify ASVs that had relative abundance patterns that were indicative of chronologic 3-month groups. Only ASVs with a maximum relative abundance ≥ 1% were considered. With this analysis we identified 14 indicator ASVs. The indicator results also supported the monthly groupings of the PCoA, as we only found indicators of month groups including February through June and August through December. No indicator ASVs were found for 3-month groups containing July or January, suggesting these are periods of transition between community types. For this reason, we described wastewater from February through June as the

spring steady state and wastewater from August to December as the fall steady state, with the primary differentiating months being February-May and August-November. We also assessed the statistical strength of these month-based community groupings with a PERMANOVA test on the Bray-Curtis distance matrix for the following groups: (1) spring = February to June, (2) fall = August to December, and (3) mix = January and July. The PERMANOVA test also supported the idea of these months having distinct bacterial communities ($R^2 = 0.212$, $p = 0.0099$).

Wastewater temperature was very tightly coupled to the change in community composition at both Milwaukee WWTPs (environmental fit, JI $R^2 = 0.96$, SS $R^2 = 0.97$; Fig. 5b) and appears to be the primary driver of the observed seasonal change in community composition. For the other measured environmental parameters, we saw differences between the WWTPs in their relationship to bacterial community composition. At SS, which receives only wastewater from a separated sewer system, all variables tested (flow rate, ammonia, total suspended solids, air temperature, phosphorus, biological oxygen demand, precipitation, and year) were significant predicators (environmental fit R^2 range = 0.35-0.85) of influent bacterial communities but were less strongly related to community change than temperature. At JI, which receives combined sewer wastewater, the environmental parameters measured were much less indicative of the community composition (environmental fit, R^2 range = 0.0035-0.28).

Sewer bacteria drive temporal trends in WWTP influent

Seasonal bacterial community variation was driven more by abundance changes of common sewer-associated ASVs than by human-associated ASVs. Dendrogram clustering of normalized

sewer-associated ASV abundances (Milwaukee time-series) illustrated that many common wastewater bacteria (e.g., *Acinetobacter*, *Arcobacter*, *Cloacibacterium*, *Flavobacterium*, *Lactococcus*) exhibited repeating temporal patterns of high/low or low/high abundance in the spring and fall states (Fig. 6). In contrast to the seasonal abundance pattern clustering of sewer-associated ASVs in the influent samples, common human ASVs exhibited less dramatic temporal fluctuations, and these changes were not predictable temporally or with the wastewater environmental data. Instead, human ASV relative abundance patterns often clustered by taxonomic affiliation (Fig. 6).

We identified two sewer-associated ASVs that exhibited significant seasonal abundance variation and one ASV matching a human fecal indicator that did not. The two sewer organisms were (1) ASV8, an indicator of September-October-November (fall-warm period) classified to the genus *Cloacibacterium*; and (2) ASV42, an indicator of February-March-April (spring-cold period) classified as *Flavobacterium*. The human fecal indicator was classified as a Bacteroides (ASV44). This ASV has 100% sequence identity to the "Human *Bacteroides* marker", a well-established marker for tracking human fecal pollution in the environment[139]. We ran autocorrelation function (ACF) with 60 1-month time lags to verify the observed seasonal relative abundance patterns of the *Cloacibacterium* and *Flavobacterium* ASVs. The autocorrelation function confirmed the repeated seasonal cycle for these two ASVs across the 5-year time series (Fig. 7a). The human specific *Bacteroides* did not show significant autocorrelations (p value = 0.05) at any time lag (Fig. 7a).

ASV-specific gene quantifications demonstrated relative abundance patterns observed in the sequence-based datasets translated to actual abundance change. The *Cloacibacterium* ASV had the highest relative and absolute abundance ranking in WWTP influent (1.1 ± 0.69%, $1.5 \bullet 10^6 \pm 1.1 \bullet 10^6$ copies/ml), followed by the *Flavobacterium* (0.41 ± 0.25%) and *Bacteroides* (0.21 ± 0.07%) ASVs (Fig. 7c). Absolute abundance quantification matched relative abundance patterns for *Cloacibacterium* (Spearman rank correlation, *rho* = 0.85; Fig. 7b), *Flavobacterium* (*rho* = 0.83) and *Bacteroides* (*rho* = 0.49). These measurements also support the observation that water temperature drives fluctuations in the resident sewer bacterial community, in that *Cloacibacterium* and *Flavobacterium* concentrations were correlated to wastewater temperature (Spearman rank correlation, *rho* = 0.90 and -0.89, respectively; $p = 2.8\text{x}10^{-9}$ and $6.4 \bullet 10^{-9}$, respectively), while *Bacteroides* concentrations were not (*rho* = -0.22, *p* = 0.30).

Milwaukee wastewater seasonality is supported spatially across the United States

Northern and southern US cities had distinct bacterial WWTP influent communities (Fig. 8). Seasonal change altered the magnitude of this regional community composition difference. For example, communities in northern cities (a cold region) during August (a high temperature period) were more similar to communities from southern US cities (a warm region) than they were to other northern communities when it was cold (Mann-Whitney U, $p < 2.2 \bullet 10^{-16}$). This similarity was greatest when southern cities experienced their coldest temperatures. Also, southern US cities, which experience less dramatic seasonal temperature change, had WWTP influent communities that were less variable than the northern cities.

Discussion

Wastewater conveyance represents a unique ecosystem in urban environments. Sewers maintain a "resident" community of microorganisms, while transient organisms continuously wash in from urban waste, runoff, and the human microbiome. In a study of a single WWTP, human gut microorganisms represented a relatively small fraction of the influent community (~7%)[129]. Our work across dozens of facilities supports this observation; we observe roughly 10-15% of the community is human-derived. The wastewater community also changes in relation to its location in the system. Human-derived microorganisms represent a larger fraction (~36%) of the community "up-the-pipe" (i.e., neighborhood sewers), but as wastewater flows through the system, resident bacteria become dominant, reaching > 85% of the assemblage. We believe this shift results from a significant increase in resident sewer bacteria, rather than a decay in human-associated microorganisms during transit. In our relatively limited testing, we found resident sewer organisms increased 2.7- to 19-fold from neighborhood sewers to the treatment plant, while human-associated bacteria stayed relatively constant (1.3-fold change).

We note that we did not attempt to partition what we term the "resident community" into organisms washing in from urban waste versus those that are truly sewer residents. Others have suggested soil bacteria may make up a significant fraction (>20%) of sewer microbes[129]. We agree that this is likely, but it is not clear if these organisms are sewer residents having originated from soil or represent transient flux into the system. Truly transient sewer organisms should have highly variable distributions in time, and nearly all of the abundant organisms in our defined resident fraction were present consistently. More work is needed to further identify the true permanent

sewer residents and the possible origins for these residents. Understanding these details would contribute to both the development of markers for sewage pollution tracking[145,146] and further the understanding of which organisms are universally present and thus likely metabolically active inside these pipes.

There are numerous places in conveyance systems that can accumulate high concentrations of actively growing sewer microbes. Biofilms attached to interior pipe surfaces represent one potentially large reservoir of resident organisms, and several studies have examined these communities (reviewed by Li et al. 2019[147]). Community compositions of sewer pipe biofilm and WWTP influent suggest there is considerable interactions between the two environments, but additional sewer habitats, such as sediments, may be contributing even larger microbial loads to the wastewater[112]. An already significant effort has been put forth to understand the products of sewer biofilm activity[147], as concrete corrosion from these activities costs more than $1 billion globally each year[148]. More work is needed in a single system to eliminate cross-system variability so that unique habitats can be identified and described.

Predominant sewer microorganisms were consistent across all the systems we examined, and also seem to be common in systems globally[112,129]. Although the same genera are present, there are stark differences in the actual bacterial composition among sewer systems, and clear diversity patterns similar to those found in other aquatic ecosystems systems like lakes or the ocean. We found that alpha diversity in WWTP influent samples remains relatively constant, but up-the-pipe, the diversity was often greater. Because human microbiome contributions were greater up-the-

pipe, individual variations in these samples and household waste streams presumably increases this diversity, but this remains to be tested. Meanwhile, we believe the large, integrated water network of conveyance systems homogenizes community inputs, obscuring rarer members prior to sampling of WWTP influent. We also found that beta diversity of influent increases as more sites or more times of the year are included, but not as more years are included. This pattern is very similar to the seasonal river, lake, or oceanic basin microbial community patterns where communities predictably cycle each year, but each system has its own unique community structure and timing of community change[22,52-54].

Our time series revealed sewer resident communities exhibit significant and repeatable temporal community change, which manifested as a seasonal cycle. This was surprising to us, as surface-water seasonal cycles such as those in temperate lakes are driven by changes in a combination of temperature and light availability, which influence primary production and ultimately start a cascade of change through the food web[152]. Sewers are below-ground and thus are buffered to large temperature changes (e.g., in Milwaukee ~8 °C difference across a year), no light is available, and there are constant exogenous nutrient inputs, so it appears much of the seasonal regime is tied to wastewater temperature change. Indeed, in Milwaukee, the seasonality of the influent wastewater bacterial community composition at both treatment plants (JI and SS) was clearly driven by wastewater temperature. Some of the other measured physical-chemical parameters also correlated to the community change (48-h precipitation, flow, ammonia, BOD, total phosphorus, TSS; Table 8), but the majority of these relationships were significant at only one of the two treatment plants (the SS plant, a separated sewer system), and the correlations were weaker than

that found for water temperature. To us, it is clear that temperature is a primary driver of bacterial community change in at least some wastewater conveyance systems.

In Milwaukee, sewer wastewater temperature change follows the change in air temperature, resulting in a roughly 3-month delay between the lowest/highest average air temperatures and the lowest/highest temperatures in the wastewater. This results in the bacterial community composition being most distinct at the wastewater temperature extremes, which occur in April (cold, ~10 °C) and October (warm, ~18 °C). Although we do not have long-term time series data from other cities, it appears water temperature plays a primary role in structuring and geographically partitioning sewer bacterial communities across vast geographical distances. Communities from northern (cold) and southern (warm) US cities were strikingly distinct, but they became more similar in comparisons of warm periods in the north to cold periods in the south. The regional warm periods in the north and cold periods in the south occur asynchronously, so there is no apparent period of community convergence across these distinct temperature regions. Also, we do not have seasonal wastewater temperatures for any southern cities, so it is unknown if the pace and timing of community change in these systems matches the two-season (warm-fall to cold-spring) setup observed in the Milwaukee dataset. It also appears that southern US cities, which have smaller air temperature ranges than most northern cities, have correspondingly less variable bacterial communities. We presume these two conditions are related, but the question of how the magnitude of wastewater temperature change impacts community composition remains to be tested.

Conclusions

Temperature dependence is clearly driving large-scale changes to the bacterial community composition in municipal wastewater conveyance systems. The temperature change results in a bacterial community that exhibits striking seasonality, but this seasonal cycle occurs in a below-ground and built/engineered system. Seasonality is more typically described for surface communities, which experience both light and temperature changes over a year. This community pattern indicates the microbial communities in built infrastructure have emergent properties comparable to the rest of the aquatic microbial biosphere; and therefore, further examination of how these microbial communities adapt to built water infrastructure is warranted. Going forward, it also needs to be determined whether temperature-driven cycles in wastewater impact engineering processes at treatment plants or alter sewer pipe corrosion rates. Wastewater treatment plant performance can vary seasonally, but it is still unclear how much of this is driven by changes in the entering community. Additionally, seasonal change in wastewater communities may represent a change in the levels of human or environmental health risk during untreated sewage release. Although the human fecal bacteria remain fairly constant temporally, the seasonal abundance shifts for common sewer organisms could be used to develop more sensitive seasonal or regional specific indicators for sewage pollution tracking. Overall, we advocate for applying microbial ecological theory developed from natural ecosystems to sewer systems. Much like in the relatively new discipline of urban ecology, there are likely theories that apply across natural and built system boundaries, but also unique paradigms that exist only in the built systems. Sewers allow for some operational control and thus could prove useful in testing theories across

boundaries, but also for understanding how urban environments alter microbial community assembly, activity, and adaptation.

Figures

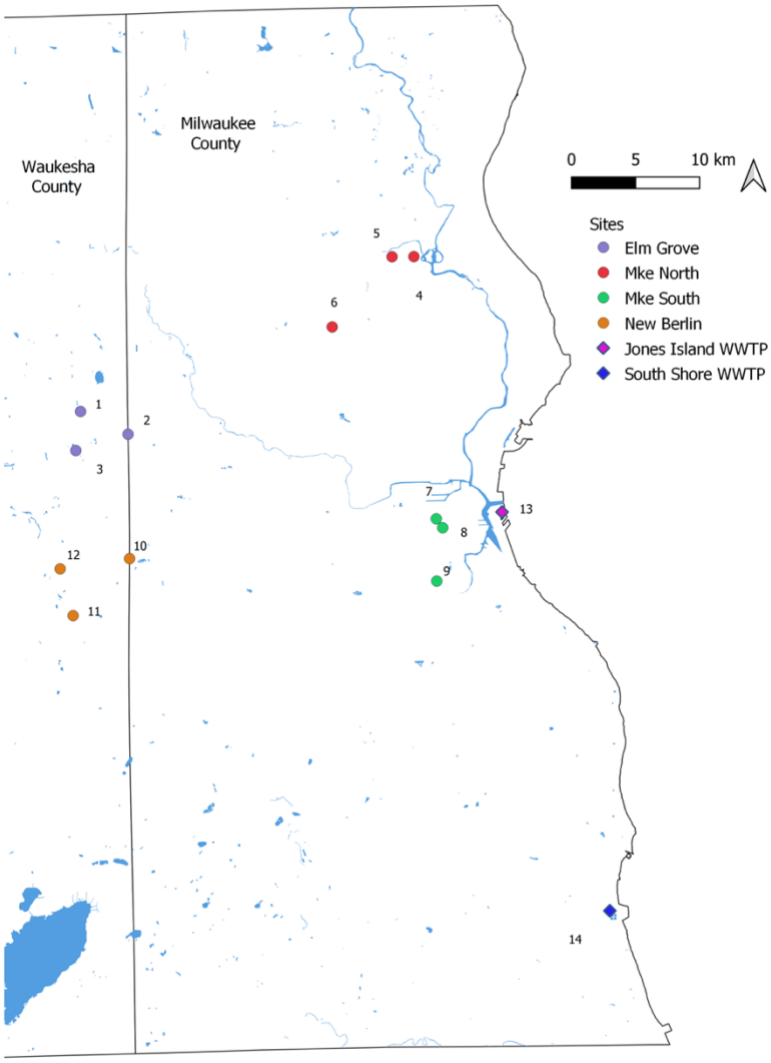

Figure 1. Map of Milwaukee sewer access.

ArcGIS skills courtesy of Emily R. Koster.

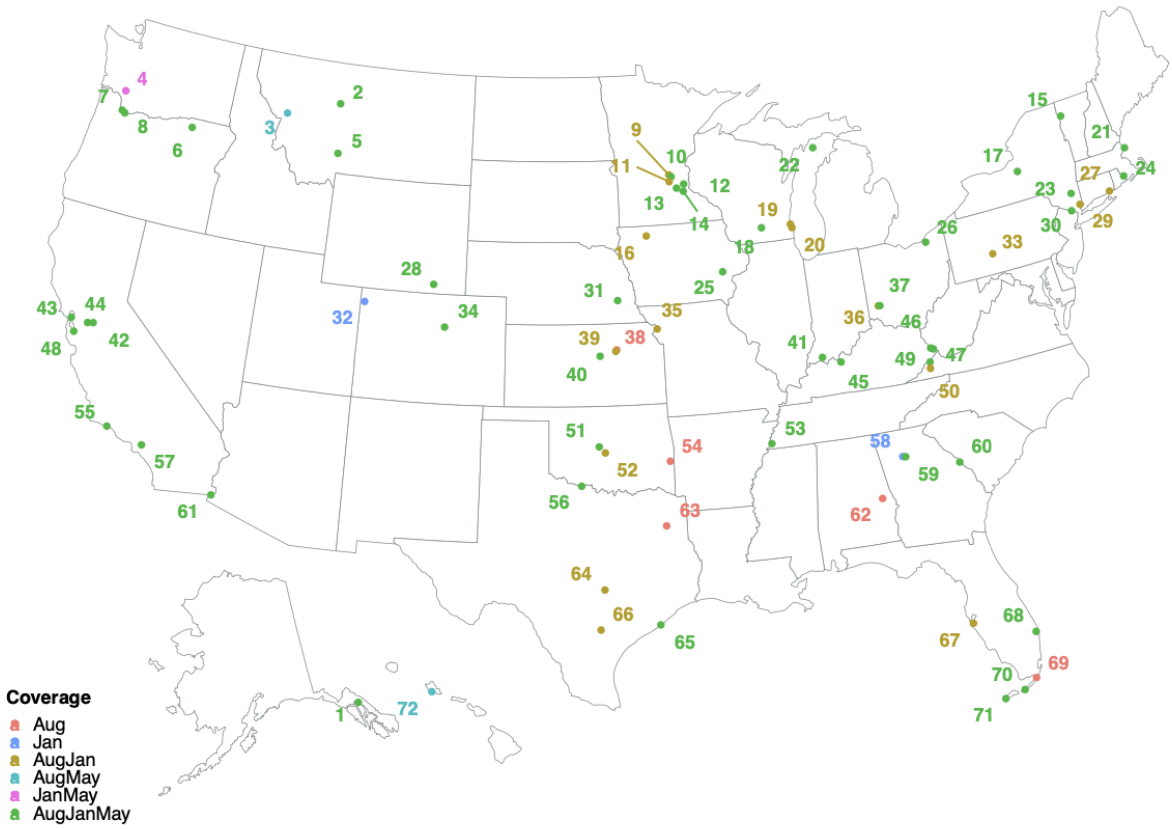

Figure 2. Map of sampled USA wastewater treatment plants.

Sampling campaign for Newton et al., 2015. WWTP influent was collected during three sampling periods designated as August, January, and May. Because influent was not collected from all three periods at each WWTP, "coverage" indicates which sampling periods that influent was collected for each location. See Table 1 for more information.

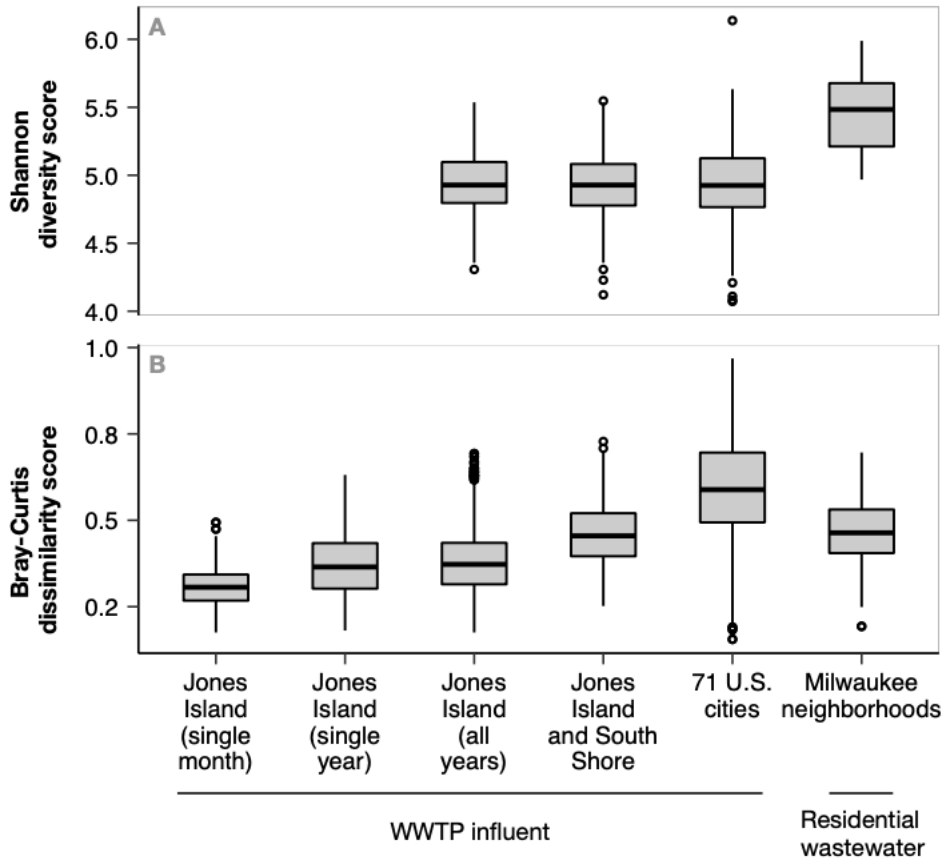

Figure 3. Community richness and dissimilarity over time and space.

(A) Alpha diversity (richness) was calculated with the Shannon diversity index and (B) beta diversity (dissimilarity) was calculated with Bray-Curtis distance metrics. Diversity was measured in wastewater treatment plant influent in the five-year time series of JI, in Milwaukee, WI; in JI and SS in Milwaukee, WI; from 77 WWTPs across the US; and in wastewater collected from residential Milwaukee neighborhood sewers. Boxes depict the median and first and third quartiles. Whisker lines extend to interquartile ranges x1.5 and points are outlier values.

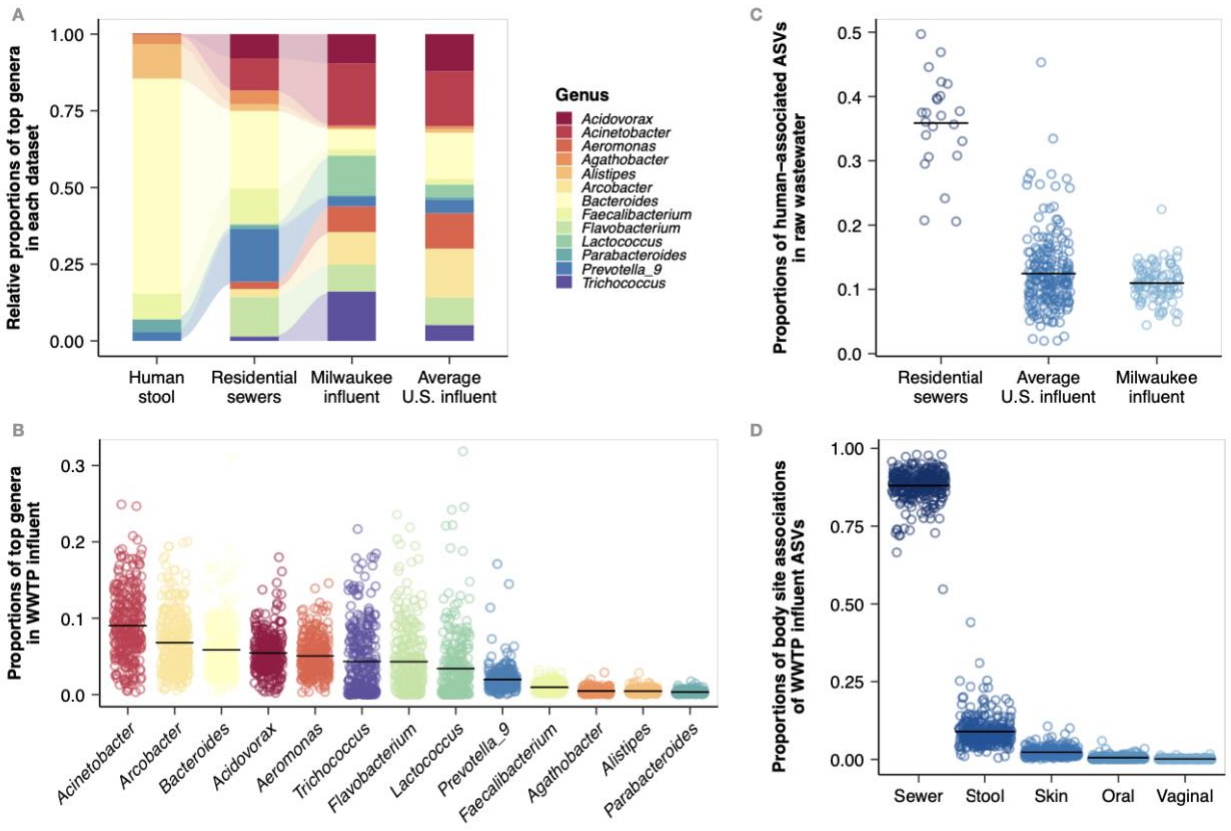

Figure 4. Microbial community changes across sources.

(A) Stacked bar plots showing the most abundant genera (top 5 from each dataset) among human stool samples from the Human Microbiome Project, wastewater from Milwaukee residential sewers, influent from a five-year time series of two Milwaukee WWTPs, and WWTPs from across the US. Bar height indicates the proportion of that genus among the abundant genera visualized. Bar colors denotes the genus. (B) Proportion of abundant genera from Figure 2A in all WWTP influent samples. (C) Proportion of all human-associated ASVs in the three wastewater datasets. (D) Proportion of ASVs from human microbiome body site sources among all wastewater samples. For B-D, circles indicate sample value and lines indicate dataset mean.

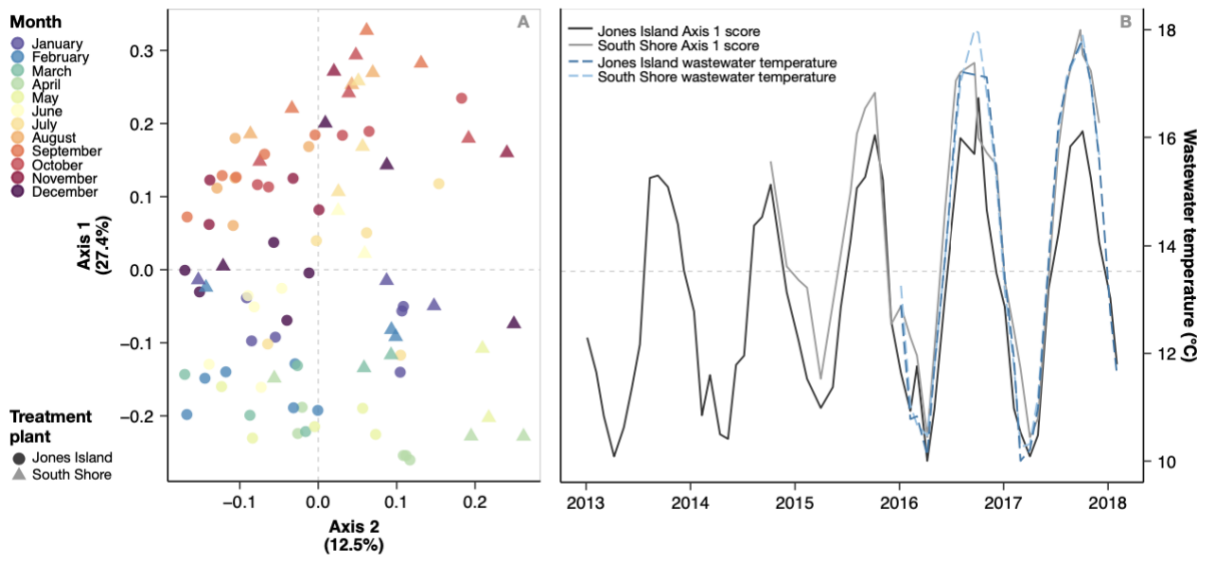

Figure 5. Microbial community dissimilarity coupled to water temperature.

(A) Principal coordinate analysis (PCoA) of influent bacterial communities from two Milwaukee WWTPs sampled once a month for five years. Points indicate influent bacterial community samples, color denotes the month sampled, and shape indicates the source WWTP. Axis 1 is set as the y-axis for visualization purposes. (B) PCoA Axis 1 scores from both WWTPs over time (solid grey lines) plotted with wastewater temperatures (blue dashed lines).

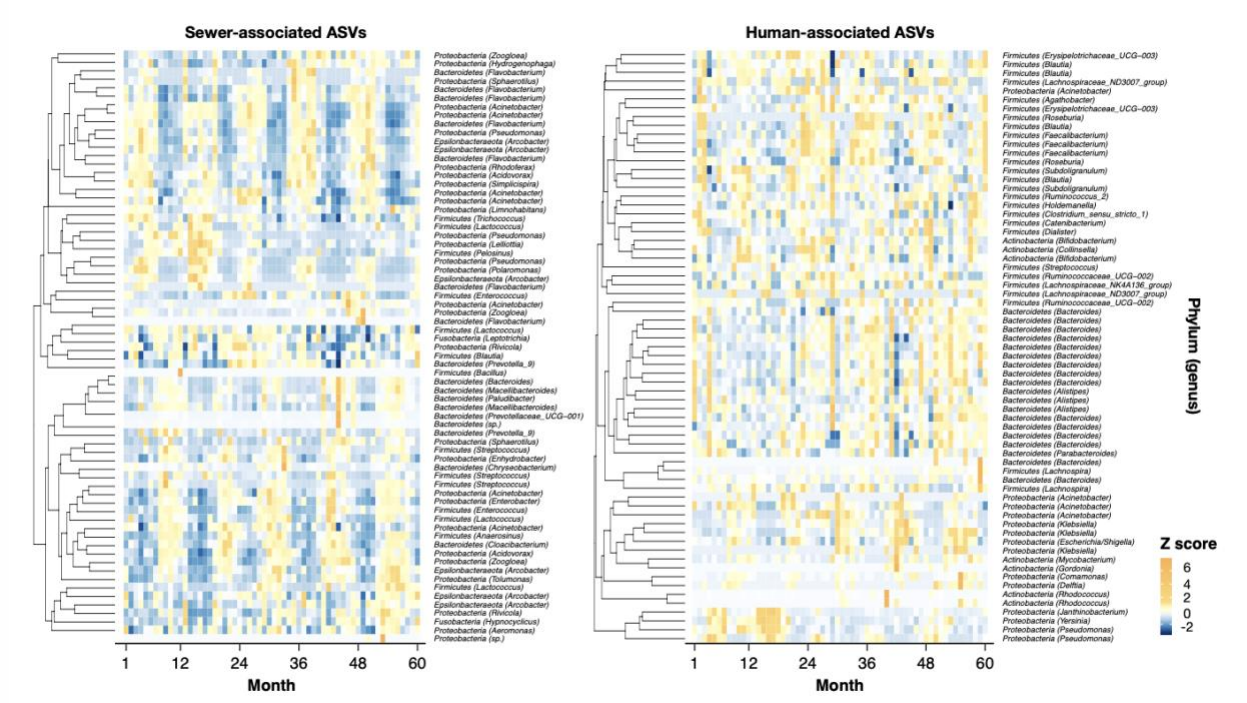

Figure 6. Common taxa clustering based on temporal abundance patterns.

Dendrograms and heatmaps of abundant (maximum relative abundance >1%) sewer-associated (left) and human-associated (right) ASVs in a five-year time series of JI influent. Heatmap colors denote within-ASV Z-scored normalized relative abundances.

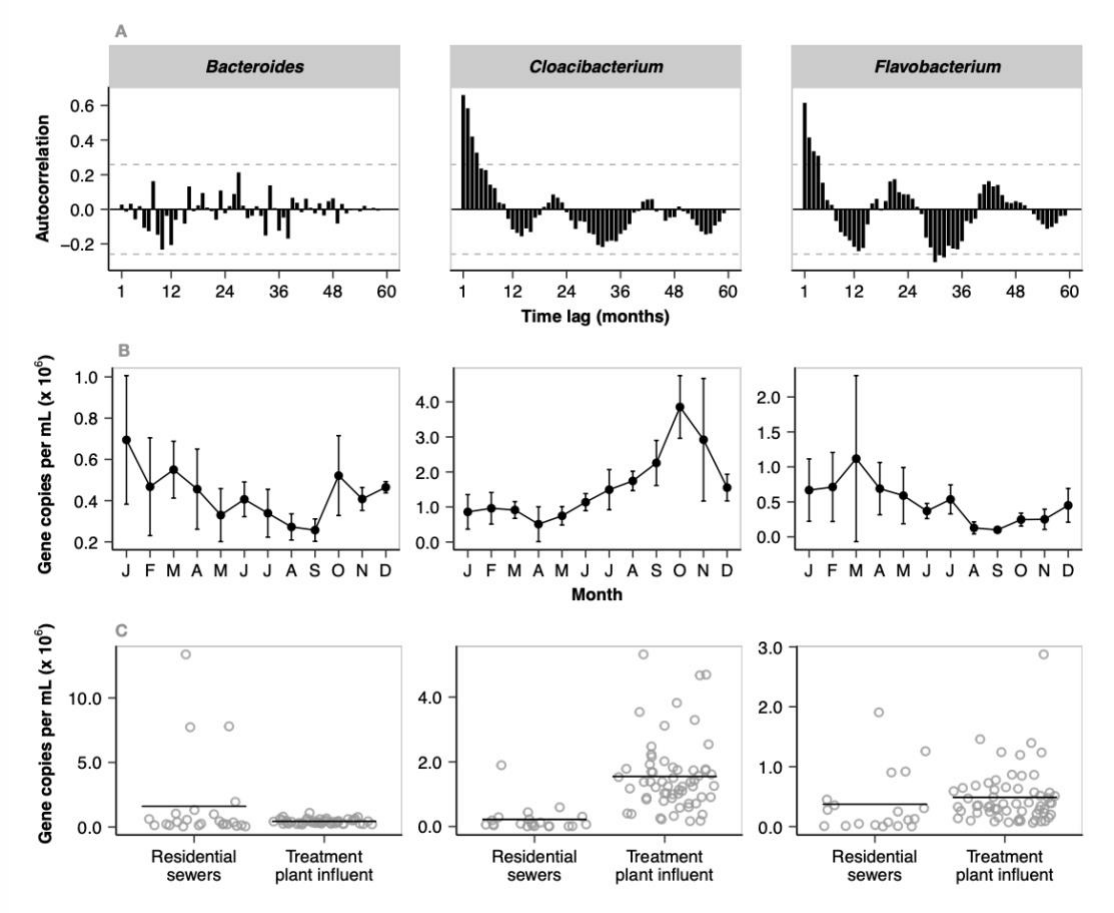

Figure 7. Time series analysis of sewer indicators and a human marker.

(A) Time-dependent autocorrelations of relative abundance change for select ASVs performed with 60 1-month time lags from the Jones Island wastewater influent dataset. Bar height indicates autocorrelation score at each time lag, and grey dashed lines indicate autocorrelation significance level (± 0.26 at $p = 0.05$). (B) Line graph of quantitative PCR measurements targeting these ASVs in JI influent. Vertical lines extend to standard deviations. (C) Quantitative PCR measurements in Milwaukee neighborhood sewers and JI influent. Horizontal lines indicate mean gene concentration. Taxonomic affiliation of the ASVs includes (left) a human-specific *Bacteroides*; (middle) a fall-associated, sewer-specific *Cloacibacterium*; and (right) a spring-associated, sewer-specific *Flavobacterium.*

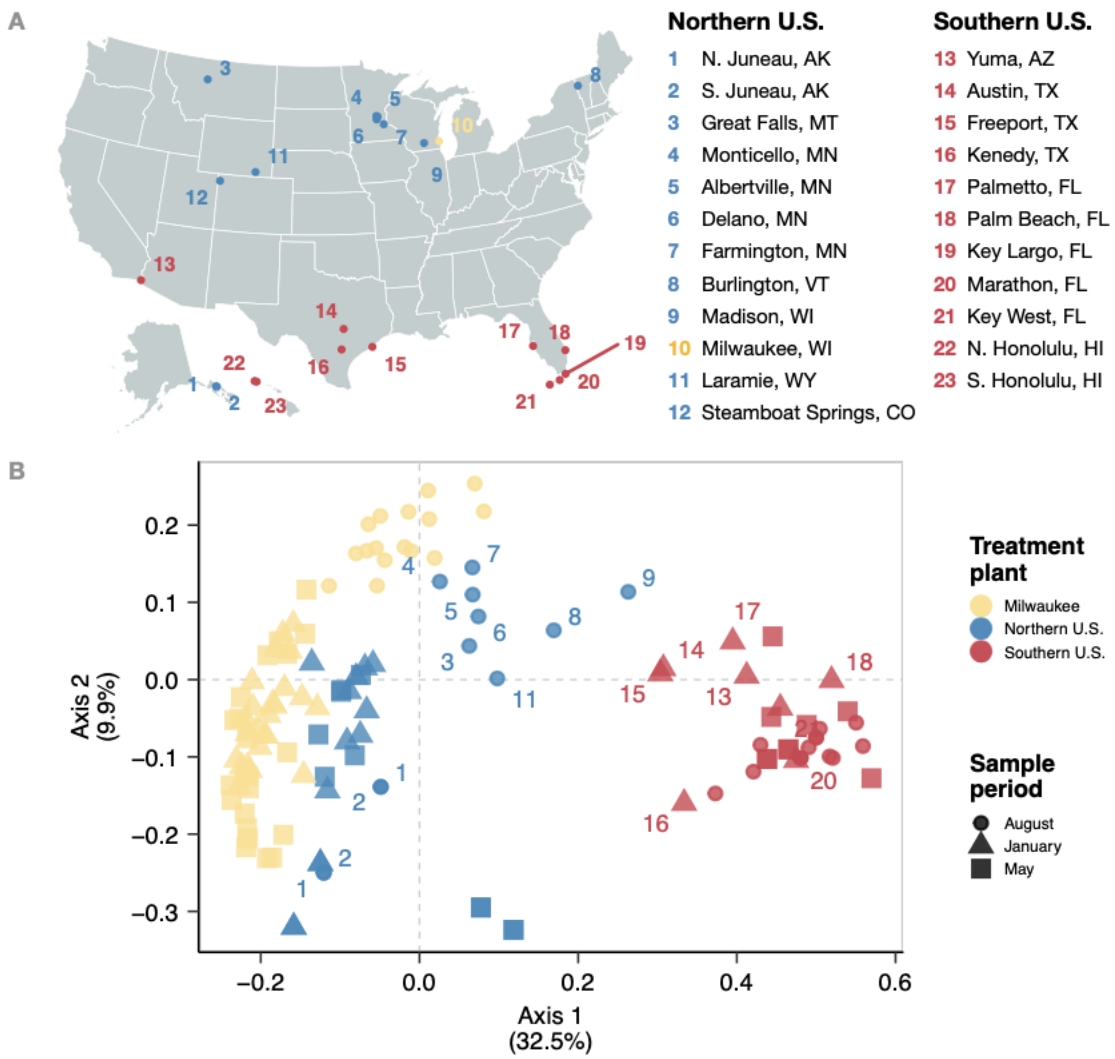

Figure 8. Wastewater community composition across the USA.

(A) map of select wastewater treatment plants sampled previously (Newton et al., 2015). (B) principal coordinate analysis of influent bacterial communities. Yellow points indicate samples from the 5-year Milwaukee WWTP time series, blue points the 11 coldest US cities in the dataset, and red points the 11 warmest US cities in the dataset. Point shapes indicate the sampling period during which wastewater was collected. Points with labels are samples from either southern US cities in January or northern US cities in August.

Tables

Table 1. Sample collection dates for Milwaukee time series.

Site	Month	Years
JI	Jan	2013, 2014, 2015, 2016, 2017, 2018
JI	Feb	2013, 2014, 2015, 2016, 2017, 2018
JI	Mar	2013, 2014, 2016, 2017
JI	Apr	2013, 2014, 2015, 2016, 2017
JI	May	2013, 2014, 2015, 2016, 2017
JI	Jun	2013, 2014, 2015, 2016, 2017
JI	Jul	2013, 2014, 2015, 2016, 2017
JI	Aug	2013, 2014, 2015, 2016, 2017
JI	Sep	2013, 2014, 2015, 2016, 2017
JI	Oct	2013, 2014, 2015, 2016, 2017
JI	Nov	2013, 2015, 2016, 2017
JI	Dec	2013, 2014, 2015, 2016, 2017
SS	Jan	2015, 2016, 2017
SS	Feb	2015, 2016, 2017
SS	Mar	2016, 2017
SS	Apr	2015, 2016, 2017
SS	May	2016, 2017
SS	Jun	2015, 2016
SS	Jul	2015, 2016, 2017
SS	Aug	2015, 2016, 2017
SS	Sep	2015, 2016, 2017
SS	Oct	2014, 2015, 2016, 2017
SS	Nov	2016, 2017
SS	Dec	2014, 2015, 2016, 2017

JI = Jones Island Water Reclamation Facility
SS = South Shore Water Reclamation Facility

Table 2. Milwaukee manhole locations and dates sampled.

Neighborhood	Manhole	Dates
Elm Grove	1	15-Dec-2015, 16-Dec-2015
Elm Grove	2	15-Dec-2015, 16-Dec-2015
Elm Grove	3	15-Dec-2015, 16-Dec-2015
New Berlin	1	15-Dec-2015, 16-Dec-2015
New Berlin	2	15-Dec-2015, 16-Dec-2015
New Berlin	3	15-Dec-2015, 16-Dec-2015
N. Milwaukee	1	15-Dec-2015, 16-Dec-2015
N. Milwaukee	2	15-Dec-2015, 16-Dec-2015
N. Milwaukee	3	15-Dec-2015, 16-Dec-2015
S. Milwaukee	1	15-Dec-2015, 17-Dec-2015
S. Milwaukee	2	15-Dec-2015, 17-Dec-2015
S. Milwaukee	3	15-Dec-2015, 17-Dec-2015

Table 3. Sample locations across USA.

Treatment plant city and state	Months and years
Juneau, AK	Jan 2013, May 2013, Sep 2012
Auburn, AL	Aug 2012
Yuma, AZ	Aug 2012, Jan 2013, May 2013
Santa Barbara, CA	Aug 2012, Jan 2013, May 2013
Palo Alto, CA	Aug 2012, Jan 2013, May 2013
Stockton, CA	Aug 2012, Jan 2013, May 2013
Discovery Bay, CA	Apr 2013, Aug 2012, Jan 2013
Whittier, CA	Aug 2012, Jan 2013, May 2013
Richmond, CA	Jan 2013, May 2013, Sep 2012
Denver, CO	Apr 2013, Aug 2012, Jan 2013
Steamboat Springs, CO	Jan 2013
New London, CT	Aug 2012, Jan 2013
West Palm Beach, FL	Aug 2012, Jan 2013, May 2013
Key Largo, FL	Aug 2012
Key West, FL	Aug 2012, Jan 2013, May 2013
Marathon, FL	Aug 2012, Feb 2013, May 2013
Palmetto, FL	Aug 2012, Jan 2013
Augusta, GA	Aug 2012, Jan 2013, May 2013
Johns Creek, GA	Aug 2012, Feb 2013, Jan 2013, May 2013
Roswell, GA	Aug 2012
Honolulu, HI	Jun 2013, Sep 2012
Iowa City, IA	Aug 2012, Jan 2013, May 2013
Spencer, IA	Aug 2012, Jan 2013
Boonville, IN	Apr 2013, Aug 2012, Jan 2013
Junction City East, KS	Aug 2012
Junction City, KS	Aug 2012, Jan 2013
Salina, KS	Jan 2013, May 2013, Sep 2012
Hardinsburg, KY	Aug 2012, Jan 2013, May 2013
Gloucester, MA	Aug 2012, Jan 2013, May 2013
Brockton, MA	Apr 2013, Aug 2012, Jan 2013

Treatment plant city and state	Months and years
Fall River, MA	Apr 2013, Aug 2012, May 2013
Delano, MN	Aug 2012, Jan 2013
Monticello, MN	Aug 2012, Jan 2013
Farmington, MN	Aug 2012, Jan 2013, May 2013
Albertville, MN	Aug 2012, Jan 2013, May 2013
Shakopee, MN	Aug 2012, Jan 2013, May 2013
St. Paul, MN	Aug 2012, Jan 2013, May 2013
St. Joseph, MO	Aug 2012, Jan 2013
Great Falls, MT	Apr 2013, Aug 2012, Jan 2013
Bozeman, MT	Aug 2012, Jan 2013, May 2013
Missoula, MT	Aug 2012, May 2013
Lincoln, NE	Aug 2012, Jan 2013, May 2013
Poughkeepsie, NY	Apr 2013, Aug 2012, Jan 2013
Syracuse, NY	Aug 2012, Jan 2013, May 2013
Bedford, NY	Aug 2012, Jan 2013
Hillburn, NY	Aug 2012, Jan 2013, May 2013
Springboro, OH	Apr 2013, Aug 2012, Jan 2013
Franklin, OH	Aug 2012, Feb 2013
Woodmere, OH	Aug 2012, Jan 2013, May 2013
Yukon, OK	Aug 2012, Jan 2013, May 2013
Heavener, OK	Aug 2012
Moore, OK	Aug 2012, Feb 2013
Portland, OR	Aug 2012, Jan 2013, May 2013
Pendleton, OR	Aug 2012, Jan 2013, May 2013
Duncansville, PA	Aug 2012, Jan 2013
Memphis, TN	Apr 2013, Aug 2012, Jan 2013
Freeport, TX	Apr 2013, Aug 2012, Jan 2013
Burkburnett, TX	Aug 2012, Jan 2013, May 2013
Austin, TX	Aug 2012, Jan 2013
Gladewater, TX	Aug 2012
Kenedy, TX	Aug 2012, Jan 2013

Treatment plant city and state	Months and years
Clintwood, VA	Aug 2012, Jan 2013, May 2013
Coeburn, VA	Aug 2012, Jan 2013
Burlington, VT	Aug 2012, Jan 2013, May 2013
Vancouver, WA	Apr 2013, Aug 2012, Jan 2013
Kelso, WA	Jan 2013, May 2013
Oak Creek, WI	Aug 2012, Jan 2013
Milwaukee, WI	Aug 2012, Jan 2013
Madison, WI	Apr 2013, Aug 2012, Jan 2013
Matewan, WV	Aug 2012, Feb 2013, May 2013
Williamson, WV	Aug 2012, Feb 2013, May 2013
Laramie, WY	Aug 2012, Jan 2013, May 2013

Table 4. *Flavobacteria* primer information.

Target	Primer name	Direction	Sequence	Anneal (°C)
Flavobacterium ASV11	flavo11_sew_629F	forward	GAACGGCCATTGATACTGCT	58-60
	flavo11_sew_859R	reverse	TAGCCACTGAAGTTGCCCCC	
Flavobacterium ASV42	flavo42_sew_630F	forward	ACGGCCATTGATACTGTCTGA	58-60
	flavo42_sew_859R	reverse	TAGCCACTGAGATTGCTCCC	
Cloacibacterium ASV08	cloaci08_sew_682F	forward	AGTGTAGCGGTGAAATGCAT	58-60
	cloaci08_sew_860R	reverse	TTGGTCTCTGAACCCTAAAGC	
Cloacibacterium ASV32	cloaci08_sew_682F	forward	AGTGTAGCGGTGAAATGCAT	58-60
	cloaci32_sew_859R	reverse	TGGTCTCTGAAGCTTGCGCT	

Table 5. *Flavobacteria* gene block positive controls.

Gene block name	16S rRNA gene V4-V5 sequence
flavo11_sew_16S	ACGGAGGATCCAAGCGTTATCCGGAATCATTGGGTTTAAAGGGTCCGTAGGCGGTTTAGTAAGTCAGTGGTGAAAGC CCATCGCTCAACGGTGGAACGGCCATTGATACTGCTAGACTTGAATTATTAGGAAGTAACTAGAATATGTAGTGTAG CGGTGAAATGCTTAGAGATTACATGGAATACCAATTGCGAAGGCAGGTTACTACTAATGGATTGACGCTGATGGACG AAAGCGTGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACTAGCTGTTGGGGGCAA CTTCAGTGGCTAAGCGAAAGTGATAAGTATCCCACCTGGGGAGTACGTTCGCAAGAATGAA
flavo42_sew_16S	ACGGAGGATCCAAGCGTTATCCGGAATCATTGGGTTTAAAGGGTCCGTAGGCGGTCAGATAAGTCAGTGGTGAAAGC CCATCGCTCAACGGTGGAACGGCCATTGATACTGTCTGACTTGAATTATTAGGAAGTAACTAGAATATGTAGTGTAG CGGTGAAATGCTTAGAGATTACATGGAATACCAATTGCGAAGGCAGGTTACTACTAATGGATTGACGCTGATGGACG AAAGCGTGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACTAGCTGTTGGGAGCAA TCTCAGTGGCTAAGCGAAAGTGATAAGTATCCCACCTGGGGAGTACGTTCGCAAGAATGAA
cloaci08_sew_16S	ACGGAGGGTGCAAGCGTTATCCGGATTTATTGGGTTTAAAGGGTCCGTAGGCGGACTTATAAGTCAGTGGTGAAAGC CTGTCGCTTAACGATAGAACTGCCATTGATACTGTAAGTCTTGAGTATATTTGAGGTAGCTGGAATAAGTAGTGTAG CGGTGAAATGCATAGATATTACTTAGAACACCAATTGCGAAGGCAGGTTACCAAGATATAACTGACGCTGAGGGACG AAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGCTAACTCGTTTTTGGGCTTTA GGGTTCAGAGACCAAGCGAAAGTGATAAGTTAGCCACCTGGGGAGTACGCTCGCAAGAGTGAA
cloaci32_sew_16S	ACGGAGGGTGCAAGCGTTATCCGGATTTATTGGGTTTAAAGGGTCCGTAGGCGGACTTATAAGTCAGTGGTGAAAGC CTGTCGCTTAACGATAGAACTGCCATTGATACTGTAAGTCTTGAGTATATTTGAGGTAGCTGGAATAAGTAGTGTAG CGGTGAAATGCATAGATATTACTTAGAACACCAATTGCGAAGGCAGGTTACCAAGATATAACTGACGCTGAGGGACG AAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGCTAACTCGTTTTTGGAGCGCA AGCTTCAGAGACCAAGCGAAAGTGATAAGTTAGCCACCTGGGGAGTACGCTCGCAAGAGTGAA

Table 6. Human microbiome body site descriptions.

Binned body site name	HMP body site description
Skin	L_Retroauricular crease
	R_Retroauricular crease
	L_Antecubital fossa
	R_Antecubital fossa
	Anterior nare
Oral	Saliva
	Tongue dorsum
	Hard palate
	Buccal mucosa
	Attached/Keratinized gingiva
	Palatine Tonsils
	Throat
	Supragingival plaque
	Subgingival plaque
Vaginal	Vaginal introitus
	Mid vagina
	Posterior fornix
Stool	Stool

HMP = Human Microbiome Project

Table 7. R packages and functions.

R package	Analysis	Function	Call	Purpose
dada2 (1.12.1)	Quality control and merge forward and reverse 16S rRNA gene amplicon sequences in FASTQ files			
phyloseq (1.28.0)	Organize ASV abundance matrices, taxonomy classifications, and sample information			
ggplot2 (3.2.1)	Create figures			
decontam (1.4.0)	Contaminant identification	isContaminant	`isContaminant(phyloseq.object, method = "prevalence", neg = "NTC")` `isContaminant(phyloseq.object, method = "prevalence", neg = "mock")`	Identify and remove reads from the no template control and mock community
vegan (2.5.6)	Shannon alpha diversity	diversity	`diversity(abundance.matrix, method = "shannon")`	Calculate alpha diversity within microbial community samples
	Bray-Curtis dissimilarity	vegdist	`vegdist(abundance.matrix, method = "bray")`	Calculate beta diversity between microbial community samples
	Euclidian dissimilarity	vegdist	`diversity(abundance.matrix, method = "euclidian")`	Calculate dissimilarity between normalized (z scores) ASV abundances
	Constrained correspondence analysis (CCA)	cca	`cca(abundance.matrix ~ ., sample.info, na.action = na.exclude)`	Ordinate microbial communities against environmental variables
	Environmental fit	envfit	`envfit(cca.result ~ ., sample.info, perm = 999, na.rm = TRUE)`	Fit environmental variables to CCA
stats (3.6.0)	Shapiro-Wilk	shapiro.test	`shapiro.test(vector)`	Test for normal distributions
	Mann-Whitney U	wilcox.test	`wilcox.test(x = x.vector, y = y.vector, paired = FALSE, alternative = "greater")`	Compare alpha diversity measurements and cluster heights between datasets
	Average hierarchical cluster	hclust	`hclust(vegdist.object, method = "average")`	Cluster dissimilarity matrices to create dendrogram of ASVs based on abundance patterns
	Centroid hierarchical cluster	hclust	`hclust(vegdist.object, method = "centroid")`	Cluster dissimilarity matrices based on distance from center of clusters

R package	Analysis	Function	Call	Purpose

	Autocorrelation	acf	`acf(vector, lag.max = 60, plot = FALSE)`
			`# extract for ggplot`
			`with(acf.result, data.frame(lag, acf))`
				Observe seasonal abundance patterns of ASVs

	Analysis of variance (ANOVA)	aov	`aov(vector ~ variables)`	Test if environmental metadata and ASV abundances are explained by season
	Spearman rank correlation	cor.test	`cor.test(vector.x, vector.y, method = "spearman")`	Correlate ASV abundances measured by 16S rRNA sequencing and ddPCR
	Model prediction	predict	`predict(aov.result, interval = "confidence")`	Make predictions from ANOVA
OTUtable (1.1.2)	Normalization	zscore	`zscore(abundance.matrix)`	Normalize ASV abundances to z scores $((x-\mu)/\sigma)$
indicspecies (1.7.6)	Multi-level pattern analysis (indicator species)	multipatt	`multipatt(abundance.matrix, variables, min.order = 3, max.order = 3, control = how(nperm = 999)`	Find associations between ASV abundances and groups of months (seasons)
ape (5.3)	Principal coordinate analysis (PCoA)	pcoa	`pcoa(vegdist.object)`	Visualize dissimilarities (as Bray-Curtis distances) between samples

CHAPTER III

WASTEWATER ANTIBIOTIC RESISTANCE GENES ARE COUPLED WITH TEMPERATURE-DRIVEN BACTERIAL COMMUNITIES

Abstract

Environments impacted by sewage are a focal point of antibiotic resistance research. Conditions in sewage promote the development and spread of antibiotic resistance genes (ARGs) and exert a selective pressure that fixes ARGs in a microbial community. Little is known about microbial communities that reside in sewers, despite them being a vast majority over allochthonous, waste-associated groups. We recently showed that microbial community assembly was strongly linked to wastewater temperature, a trend that was consistent over years within a single city and echoed across the United States. However, it is not clear whether predictable community dynamics would translate to trends in ARG abundance and diversity. Here, we used metagenomic sequencing to catalog all possible ARGs in wastewater treatment plant (WWTP) influent collected periodically over five years from six US cities. We supplemented compositional changes of ARGs in metagenomes with absolute quantifications of common sewage ARGs (*sulI*, *qnrS*, *tet(A)*, *tetO*) and *intI1* using droplet digital PCR (ddPCR). Metagenomics revealed a strong correlation of wastewater ARG composition to the source city and sample month. Cities had distinct catalogs of ARGs, with northern- and southern-most cities having the highest number of indicators. ARGs were present in wastewater year-round but fluctuated in relative abundance with seasonal changes. In other words, wastewater from a city may have its own pool of ARGs, and those ARGs increase or decrease with the bacterial host response to wastewater temperature. Our results highlight the connectedness of ARGs to the sewer ecosystem, such that potentially both microbial community dynamics and the ecology of each system impacts the composition of ARGs in sewage.

Introduction

The current arsenal of antibiotics that effectively treat persistent infections has been rapidly diminishing for decades. Rampant antibiotic resistance is a consequence of negligent antibiotic consumption since its discovery a century ago. Antibiotic exposure puts a selective pressure on antibiotic resistance genes (ARGs) that favors the survival of resistant bacteria. Selective pressure is exerted on gut microbiota and any other environment where antibiotics are found. Up to 80% of ingested drugs are excreted in urine[153] so environments polluted with sewage or animal waste have potential to exert selective pressure on ARGs. Urban sewers collect waste from a variety of sources, such as residences, industries, and hospitals, creating complex mixtures of microbes, pharmaceuticals, and chemicals that arguably have never occurred on earth before and may never be repeated again. Microbial communities with permanent residence in sewers have adapted means to survive this stressful environment, including carrying ARGs[154–158].

Knowledge gaps in environmentally derived resistance lie in part with the limited understanding of the sewer microbiome. Previous work has shown that sewage communities maintain relatively consistent structure[4,100] and assembly was strongly dictated by wastewater temperature[159]. However, it has yet to be seen how community-level dynamics influence wastewater ARG composition and distribution. Here, we examined the composition of ARGs between cities and over time to capture ARG relationships to community changes. We sequenced metagenomes from 30 untreated wastewater samples, covering six USA cities over six years. With droplet digital PCR (ddPCR), we quantified five ARGs (*sulI, tet(A)*, *tetO*, and *qnrS*) and the class 1 integron-integrase

gene (*intI1*) in a five-year wastewater time series. The same genes were quantified in wastewater from 12 US cities collected three times in one year, and *intI1* quantified in an additional 39 cities.

Material and Methods

Sample information

Wastewater treatment plant (WWTP) influent was collected from two WWTPs (Jones Island and South Shore) in Milwaukee, WI, USA once a month from January 2013 to December 2017, and twice a month in April, May, September and October 2018, using methods described previously[159]. WWTP influent was collected from Palo Alto, CA; Key West, FL; Laramie, WY; and Gloucester, MA, USA in Aug 2012, Jan 2013, and May 2013, described previously[4]. In October and November 2018, twice each month, WWTP influent was collected from the Sand Island, HI, USA WWTP. Sample location and dates can be found in Tables 8 and 9.

Metagenomic sequencing and data processing

Metagenomic shotgun sequencing was performed by the Ramaciotti Centre for Genomics (Sydney, Australia), using Illumina HiSeq 2500 sequencing technology and Nextera XT kit (Illumina) for library preparation (2x150 bp). Low-quality reads were removed from FASTQs using parameters established by Minoche, et al.[160], with iu-filter-quality-minoche v.2.12 from the package suite illumina-utils[161]. Filtered paired-end reads were assembled with MEGAHIT 1.2.9[162] and contigs less than 300 bp were removed using SeqKit v.2.2.0[163].

Antibiotic resistance gene mining

Contigs were aligned to the Comprehensive Antibiotic Resistance Gene Database (CARD v.3.2.5) using the tool Resistance Gene Identifier (RGI main v.5.1.1)[164]. Only "perfect" (exact match to reference sequence) and "strict" (previously unknown variant detected by model) alignments were analyzed. A custom R[141] script converted RGI alignment files to an ARG abundance table by counting the number of instances an Antibiotic Resistance Ontology (ARO; https://www.ebi.ac.uk/ols/ontologies/aro) ID aligned to contigs in each sample. ARG counts were normalized to reads per kilobase million (RPKM):

$$RPKM = \frac{n}{N * 10^6 * l}$$

Where:

$n =$ number of times an ARG aligned to the sequence file (a cell in a sample-by-ARO ID abundance matrix)

$N =$ sum of total number of times an ARG aligned to the sequence file (the sum of a sample in the abundance matrix)

$l =$ length of the reference gene (multiply the number of characters in the "CARD protein sequence" column of the RGI alignment file by three) in kilobase pairs (divide by 1000)

Gene quantification

Antibiotic resistance genes *tet(A)*[165], *tetO*[166], *sulI*[158], *qnrS*[158], and the class 1 integron-integrase gene *intI1*[167] were quantified in WWTP influent samples using Bio-Rad QX200 Droplet Digital PCR (ddPCR) system. Duplicate reactions were set up in 96-well plates as follows: 11-µl EvaGreen Supermix (Bio-Rad, #1864034), 1.3-µl 5-µM forward and reverse primers, 6.4-µl sterile water, and 2-µl 100x-diluted DNA template. PCR was run on a vapo-protect Mastercycler pro S

under the following conditions: 95°C for 5 min; 40 cycles of 95°C for 30s, target-specific annealing temperature (Table 10) for 1 min; 4°C for 5 min; 90°C for 5 min; 4°C hold.

Results were downloaded to CSV files from the plate reader program QuantaSoft (Bio-Rad, #12012172) and organized in R[141]. Duplicated assays with variation coefficients (CV) greater than 0.5 were re-run and/or discarded. The mean copies-per-20-µl-well measurement was converted to ARG copies-per-ml of the original wastewater sample:

$$ARG_{copies/ml} = \frac{\mu * d * e}{v * f}$$

Where:

$\mu =$ mean copies per 20-µl well of duplicate assays
$d =$ dilution factor of DNA template
$e =$ elution volume of extracted DNA (µl)
$v =$ volume of DNA template in PCR (µl)
$f =$ volume of original sample fixed to filter (ml)

Statistics and graphics

DNA sequence analysis

All graphics were created using the R package ggplot2[144]. Bray-Curtis distances between samples of ARG RPKM values were calculated using *vegdist* from the R package vegan v.2.6.2[168]. A principal coordinate analysis (PCoA) was performed on distance matrices using the *pcoa* function from ape 5.6.2[143]. Dendrograms (ggendro v.0.1.23[169]) were created from hierarchical clustering ("average" agglomeration method; R stats[141] package) Bray-Curtis distances between samples and ARGs.

Shapiro-Wilk tests for normality (R stats[141] function *shapiro.test*) were run prior to any comparison statistics. Bray-Curtis distances were fed to a goodness of fit test (function *envfit*) to assess the significance of month sampled and WWTP location on ARG *composition* (RPKM values). Month and location were similarly tested on ARG *content* (presence/absence), substituting Jaccard distances between binary samples. To test if month or location yielded the most dissimilarity in the dataset, samples collected from (A) different locations during the same month and (B) the same location over multiple months were compared. Bray-Curtis distances between samples within groups A and B were compared using a one-sided Wilcoxon rank-sum test (R stats[141] function *wilcox.test*).

Gene quantification

The relative change from the mean of ARGs (i.e., over time in Milwaukee dataset, between locations in USA dataset) was calculated by dividing the difference of an observation (n) and the mean (μ) by the mean:

$$relative\ change = \frac{n - \mu}{\mu}$$

Where:

$n =$ ARG copies per ml original sample
$\mu =$ mean copies per ml in that sample group
fold change of 1 = an observation increased 1-fold (2-times) from the mean

A permutational multivariate analysis of variance (PERMANOVA, vegan[168] function *adonis*) measured the impact of month sampled on concentrations for each ARG. One-sided Wilcoxon

rank-sum tests (R stats[141], *wilcox.test*) were run to assess which sample groups (i.e., month, north, south) had higher ARG concentrations. Thiel-Sen median-based linear regressions (function and package *mblm*) measured the correlation of *intI1* concentrations to the other ARGs.

Results

Wastewater ARG composition wastewater is influenced by geography and season

ARG composition in wastewater samples (resistomes) were compared using the Bray-Curtis dissimilarity metric, calculated from ARG abundances normalized to reads per kilobase million (RPKM). Bray-Curtis scores were then ordinated in a principal coordinate analysis (PCoA, Fig. 9). WWTP location and sampling month strongly explained resistome dissimilarity (goodness of fit, $R^2 = 0.85$, $p = 0.001$; $R^2 = 0.44$, $p = 0.003$; respectively). For WWTPs sampled in January, May, and August (CA, MA, WI, WY, FL), August samples were most distant from the other months, excluding FL, which had little variability among samples. Resistomes from FL, the second-most southern sampling site, were closest to those from HI, the southern-most site. Resistomes from more northern sites were most similar to FL and HI during warmer sampling periods (i.e., WY and SS in August).

WWTP location explained the most variability in the dataset (Wilcoxon rank sum test; $p = 6.8 \cdot 10^{-9}$; Fig. 10). To compare its influence on time sampled, we measured resistome variability (A) at the same WWTP between months, and (B) during the same month between WWTPs. There was significantly more variability when comparing different WWTPs than different months (Bray-Curtis dissimilarity, $\mu(B) = 0.41 \pm 0.12$, $\mu(A) = 0.31 \pm 0.087$, respectively). The greatest

dissimilarity (0.66) was between CA (Palo Alto, CA) and SS (Milwaukee, WI) in August. Within a single WWTP, the greatest dissimilarity (0.47) was at SS between May and January. The most similar samples were from HI (Sand Island, HI), where in October it was most similar to itself in November (0.14), and again to itself in two October samples (0.15).

Seasonal fluctuations of ARGs created within-WWTP variability (Fig. 11). Between WWTPs, there were significant changes in the presence/absence of ARGs (binary goodness of fit, $p = 0.002$) but not months ($p = 0.91$). In other words, the pool of ARGs found at a given WWTP remains consistent over time, but abundances of ARGs fluctuate, and ARG pools vary between locations. This was most evident when comparing FL and WY samples. ARGs with the highest RPKM in WY (e.g., *IND-14*, *AAC(6')-Ii*, *aadA27*) were negligible in FL (Fig. 11).

The 20 most common drug classes to which ARGs were predicted to confer resistance are shown in Fig. 12. On average, the three most abundant drug classes were aminoglycosides (12.9 ± 1.5%), macrolides (11.2 ± 3.0%), and tetracyclines (9.84 ± 1.8%). Out of 124 unique classes, ARGs conferring resistance to a single drug were the most common (22.1%), followed closely by two drugs (20.6%) and three drugs (16.0%). The ARG conferring resistance to the greatest number of drugs was *oprM*, an efflux pump for 16 different antibiotics, and was present in all 30 samples. Efflux was a common mechanism of ARGs (8.2%) but was second to antibiotic inactivation (78.4%).

An indicator species analysis identifies species that significantly correlate to groups of samples. In this case, the analysis was used to identify what and how many ARGs were unique to WWTPs. Samples from August, January, and May of 2012-2013 were put into the analysis, as six WWTPs were evenly surveyed during these periods (Table 8). Out of 3162 possible ARGs, 469 were classified as significant ($p < 0.05$) indicators of WWTPs. FL, WY, and JI had the highest number of indicators (157, 100, and 78, respectively) and SS, MA, and CA had the lowest (54, 50, and 30, respectively). The most common drug classes of indicator ARGs conferred multidrug-resistance to carbapenems, cephalosporins, and penams (85, 18%); aminoglycosides (8.1%); and cephalosporins and tetracyclines (both 4.3%). All ARGs in the carbapenem-cephalosporin-penam group encoded beta-lactamases from groups *OXA* (64), *GES* (10), *SHV* (8), *GOB* (2) and *PNGM* (1).

Wastewater ARG abundances fluctuate with warm-dominating bacterial communities

ARGs and *intI1* were quantified using ddPCR in wastewater samples collected in January, April, July, and October from 2013 to 2017 in Milwaukee, WI (Fig. 13). Results showed that ARGs concentrations were consistently highest in October (one-sided Wilcoxon rank sum test, $p = 0.005$) and lowest in April ($p = 0.002$). ARG concentrations were also closely coupled to each other and could all be predicted by *intI1* concentration (MBLM, *sulI* $p = 9.6 \cdot 10^{-5}$, *qnrS* $p = 2.1 \cdot 10^{-4}$, *tet(A)* $p = 2.1 \cdot 10^{-4}$, *tetO* $p = 1.3 \cdot 10^{-3}$). Sample month was also predictive of ARG concentrations (PERMANOVA, $p < 0.001$).

ARGs were quantified in 12 cities across the USA sampled in August 2012, January 2013, and May 2013 (Fig. 14). Latitude was not a good predictor of ARG concentrations overall (PERMANOVA, $p = 0.143$) but there were some gene-specific consistencies. *Tet(A)* was highest in wastewater from low-latitude (< 40°) cities (MBLM, $p = 0.02$) while *tetO* was greatest at high latitudes (> 40°) (MBLM, $p = 0.003$). Also, Bedford, NY had the least variable differences in fold-change (CV = 0.14) and Freeport, TX most variable (CV = 1.1). Juneau, AK overall had the highest fold-changes for three out of five ARGs (*intI1*, *sulI*, *tetO*) while Bedford, NY had the lowest for four out of five ARGs (*tet(A)*, *tetO*, *sulI*, *qnrS*).

The class 1 integron-integrase gene, *intI1*, was quantified in 41 cities across the USA (Fig. 15 and Table 9). For ease, WWTP locations were split into groups "north" and "south" depending on if their latitude was above or below 40°N, respectively. Sample period (August, January, or May) and latitude group (north or south) correlated to *intI1* concentrations (PERMANOVA, $p = 0.001$ and 0.004, respectively). In the south (latitude < 40°N), *intI1* concentrations were higher than the north (one-sided Wilcoxon rank sum test, $p = 0.001$). May samples had the highest *intI1* concentrations ($p = 4.1 \cdot 10^{-4}$) and January samples had the lowest ($p = 7.7 \cdot 10^{-6}$).

Discussion

Despite global intervention, infections resistant to the current arsenal of antibiotics are escalating. Monitoring the issue has shown not only the positive impacts of intervention strategies[170] but also temporal dynamics in antibiotic resistance. Temporal studies of resistance commonly analyze antibiotic prescriptions[171], infections[172], and how they are connected[170,173–175], but do not include

environmental monitoring of antibiotic resistance genes (ARGs) or bacteria. There has been some interest in time series analysis of wastewater treatment plants (WWTPs) that cultured resistant bacteria[176,177], quantified ARGs[156,177], and related them to physiochemical measurements[176,177]. Some environmental studies also included prescription rates[156] or measured antibiotic residues in the wastewater[176]. More rarely do they consider microbial community composition[176] or natural waterways[178]. WWTP time series are often limited to one location or short time scales[176,177], which masks if and how any observed trends in resistance are consistent. A meta-analysis showed seasonal patterns in ARGs and resistant isolates at WWTPs[179], but authors explained this with the fact that human illnesses increase in winter.

The focus on clinically relevant bacteria commonly seen in environmental studies ignores the potential influence of the native community. Further, culture-based methods are estimated to capture less than 1% of microbial diversity[180] and neglect the likelihood that microbes depend on their environment and/or community functions to grow[181]. Previous work showed microbial community assembly in wastewater is strongly dictated by temperature[159], and the study described here showed ARGs may fluctuate seasonally with communities. It is not clear if this is because (a) the majority of ARGs are carried by environmental bacteria that fluctuate seasonally; (b) microbes that carry ARGs are influenced by the community through grazing, competition, etc.; (c) seasonal changes in waste sources, such as antibiotic consumption and illness in humans and animals; and/or (d) some combination of these or other unknown influences.

This study revealed the potential that wastewater ARGs are modulated by temperature, and perhaps proliferate with the bloom of warm-dominating microbial communities. It showed that abundances of ARGs in wastewater fluctuated with the seasonal assembly of microbial communities (Fig. 9). Wastewater collected from the southern- and northern-most locations had the highest number of ARGs that were not found in any other samples. Many ARGs that were unique to certain locations were clinically relevant. The most common encoded beta-lactamases that conferred multidrug resistance to last-resort antibiotics, and were considered high risk in a recent assessment[182]. Therefore, spatial metagenomic analyses of wastewater ARGs could reveal novel ARGs that have not spread outside its source region.

It is possible that extremes in climate, being very hot, cold, or variable, may select for certain ARGs or the mobile genetic elements (MGEs) or bacteria that carry them. Given this analysis did not include the temperature of the wastewater sample, this observation is limited. Temperature was known in samples from a Milwaukee time series that had gene targets (*intI1*, *tet(A)*, *tetO*, *sulI*, *qnrS*) quantified with droplet digital PCR (ddPCR, Fig. 13). This method showed a positive correlation between ARG concentrations and temperature that was consistent over five years. Horizontal gene transfer rates have been shown to increase with temperature[183], so it is possible this is causing ARGs to proliferate in warm wastewater.

The influence of temperature was less obvious when quantifying ARGs from wastewater across the US. ARGs had no temperature correlation (Fig. 14), but the class 1 integron-integrase gene, *intI1*, showed both geographic and seasonal trends (Fig. 15). Concentrations of *intI1* were highest

in samples from southern latitudes (< 40°N) and in samples collected in May, versus January or August. The intI1 assay was done on more samples than ARGs, so it is possible that geographic and/or seasonal patterns of ARGs could be revealed with more samples. Another reason for discrepancies between the Milwaukee samples and those from other cities could be due to user handling; Milwaukee samples were collected and processed by the same people, whereas different WWTPs had their own people collecting and shipping samples to Milwaukee.

Conclusions

This study revealed the potential that wastewater ARGs are modulated by temperature, and perhaps proliferate with the bloom of warm-dominating microbial communities. It highlights insight that can be gained when incorporating microbial community analysis to longitudinal monitoring of antibiotic resistance for both culture- and molecular-based studies. This knowledge can aid research that studies the evolution of antibiotic resistance by showing potential links between taxa, environmental conditions, and ARGs. It can also enhance our understanding of mitigation outcomes by considering ecosystem dynamics that promote or curb the spread of resistance.

Figures

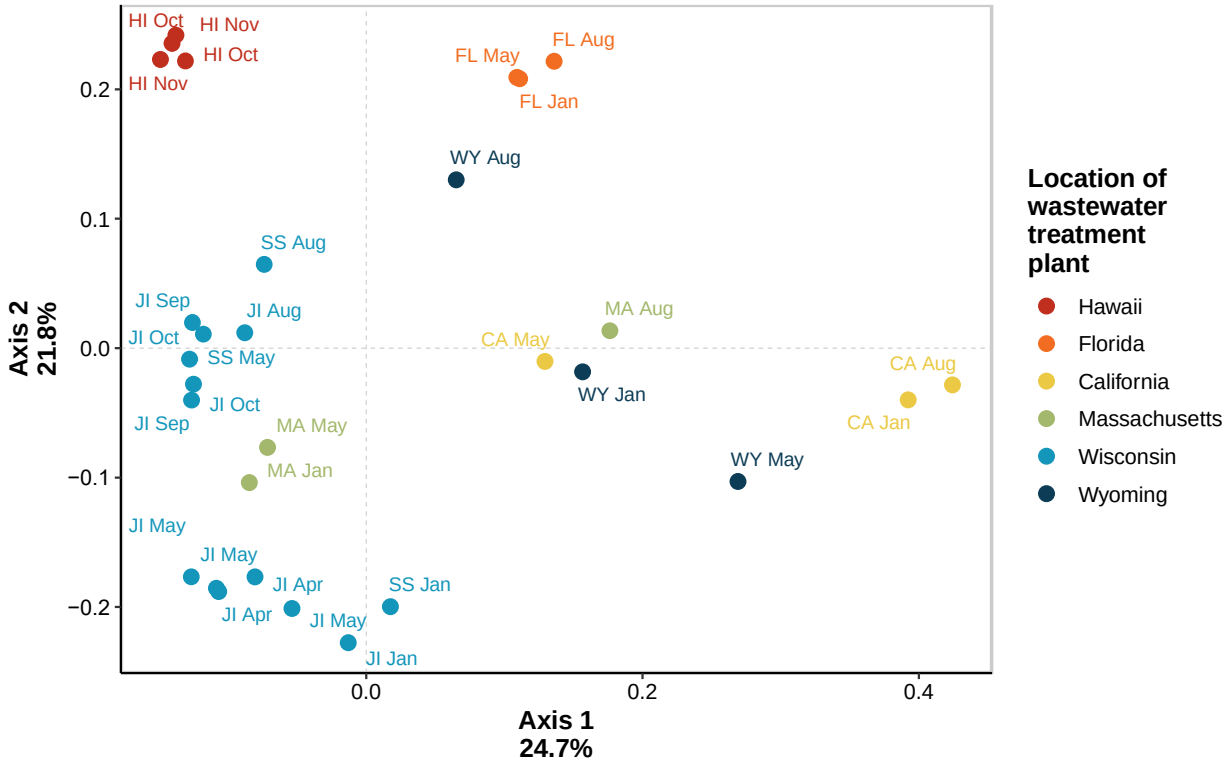

Figure 9. Wastewater resistomes across locations and seasons.

Principal coordinate analysis (PCoA) of antibiotic resistance gene profiles mined from raw wastewater metagenomes. Colors of points indicate the location of the wastewater treatment plant sampled: red: Sand Island, HI; orange: Key West, FL; yellow: Palo Alto, CA; green: Gloucester, MA; light blue: Milwaukee, WI; dark blue: Laramie, WY. Point labels denote the wastewater treatment plant short name (state abbreviations for all except Milwaukee, which had two locations, Jones Island (JI) and South Shore (SS) and month of sampling.

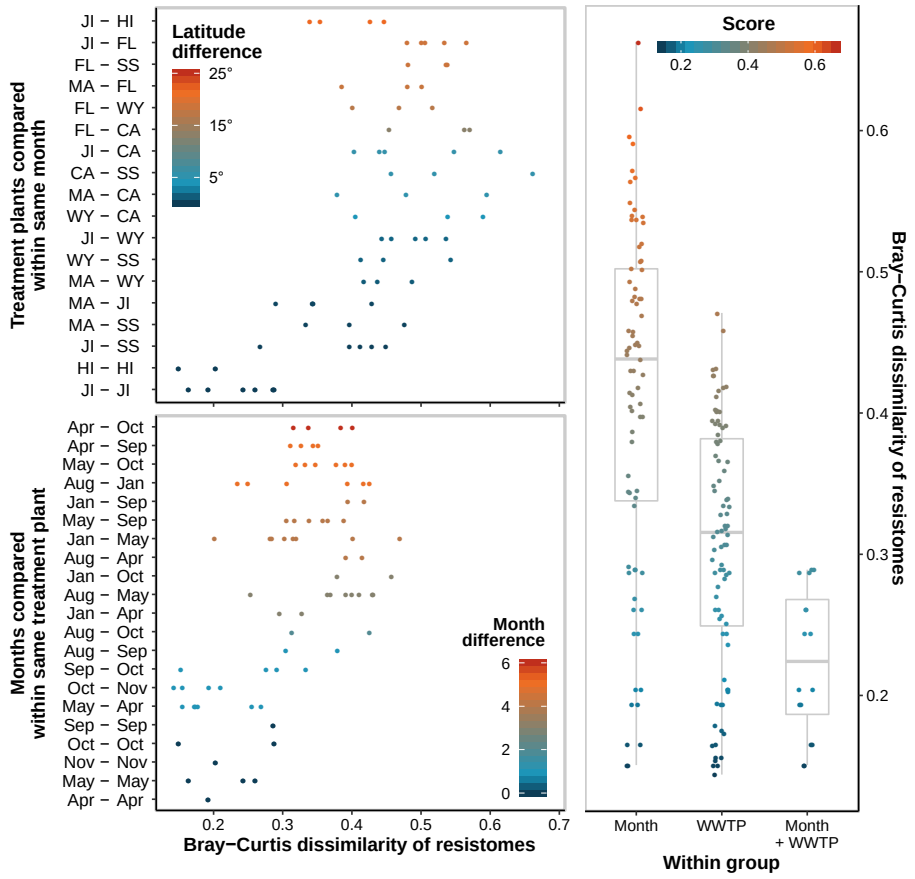

Figure 10. Wastewater resistome dissimilarity within treatment plants and months.

(top) Bray-Curtis distances comparing ARGs in wastewater collected in the same month. Each point is a distance value between two samples, and colors of points show the difference in latitude between the two treatment plants (minimum = 0°, same treatment plant; maximum = 22°, between Milwaukee, WI and Sand Island, HI). (bottom) Bray-Curtis distances comparing ARGs in wastewater collected at the same treatment plant at different times of year. Each point is a distance value between two samples, and colors of points show the cyclical difference between the months sampled (minimum = 0, same month; maximum = 6, between April and October). Samples from different years were compared, such that samples from April 2013 and November 2018 were different by 5 months, for instance. (right) Boxplots summarizing left two plots, comparing distances between wastewater ARG samples from (Month) different treatment plants, within the same month; (WWTP) different months, within the same treatment plant; and (Month + WWTP) the same month and treatment plant.

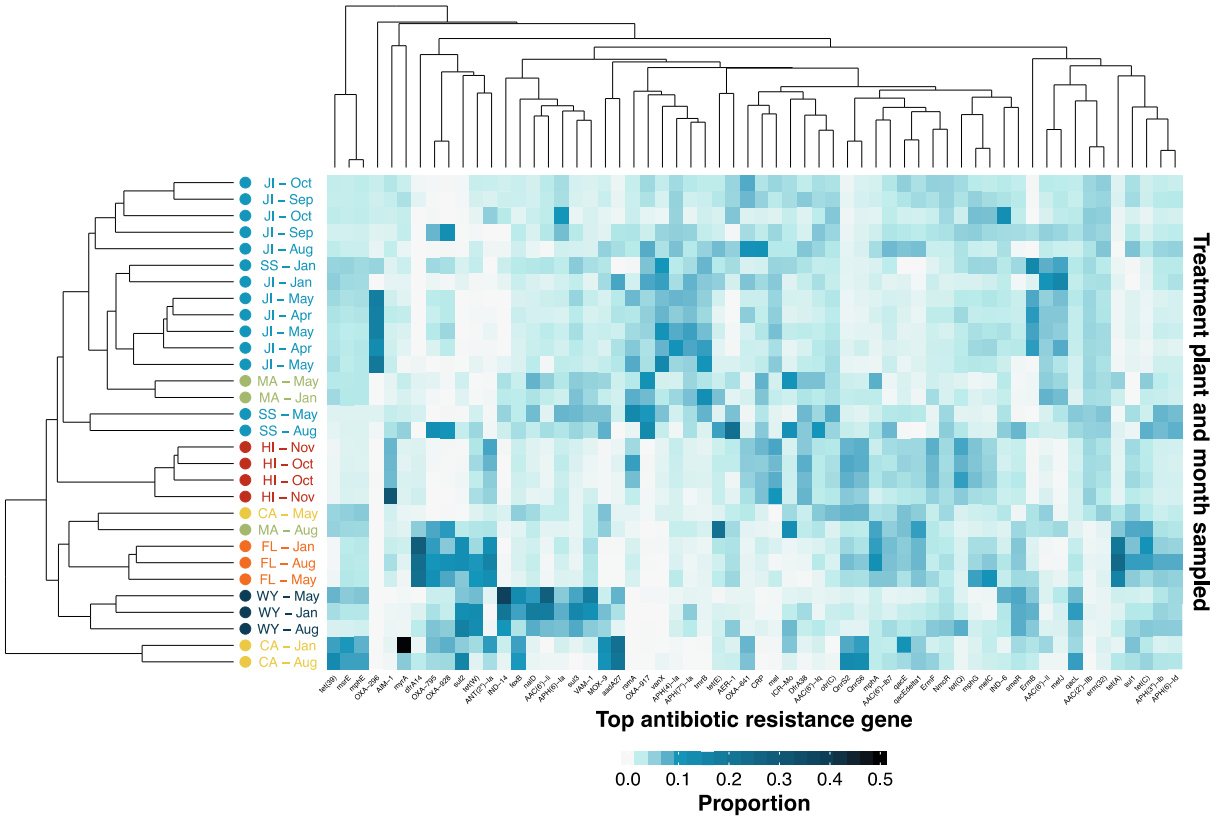

Figure 11. Resistomes clustered by geographic abundance patterns.

(left) Dendrogram of raw wastewater samples, created from hierarchical clustering of Bray-Curtis distance of RPKM values for all ARGs. (top) Dendrogram of the 60 most abundant ARGs, determined by obtaining the 25 most abundant ARG from each WWTP. (center) Heatmap of most abundant ARGs, normalized to relative abundance such that the sum of an ARG's RPKM equals 1. Heatmap ordered on the x-axis by the ARG dendrogram and y-axis by wastewater sample dendrogram.

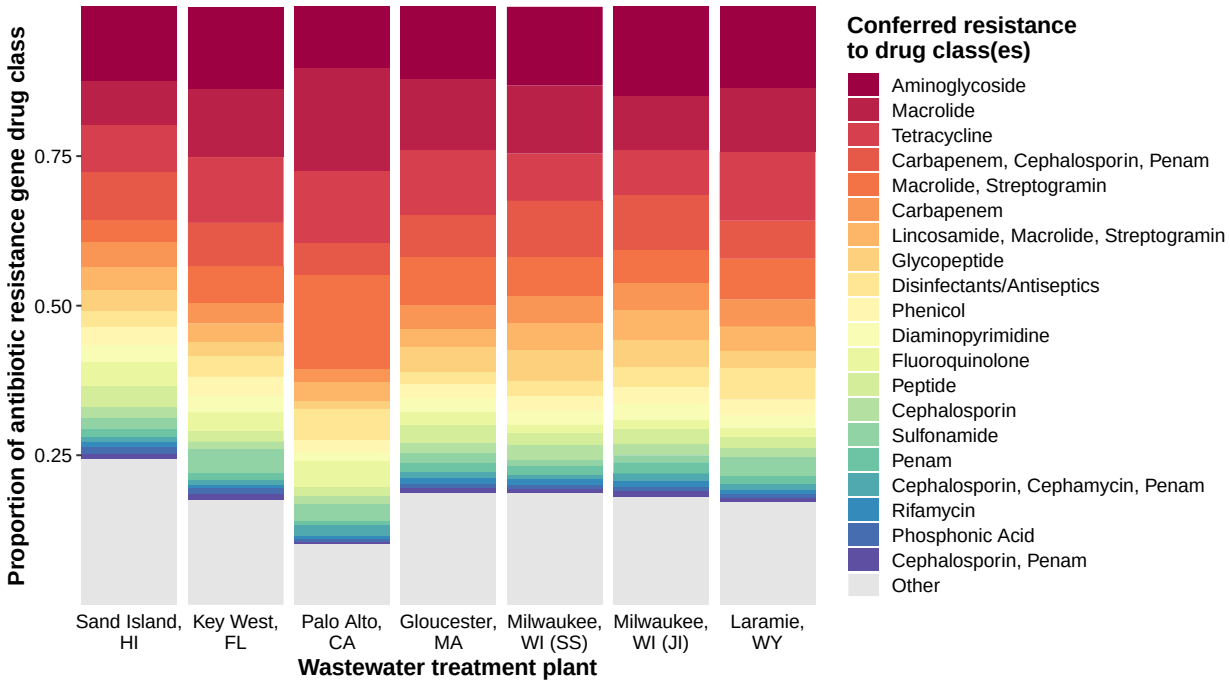

Figure 12. Drug classes of wastewater antibiotic resistance genes.

Stacked bars showing proportions of drug classes to which ARGs confer resistance. Colors of bars show ARG drug class(es), and height of bars show the proportion of drug classes in each state. Drug classes that were not in the 20 most abundant were categorized as “other” and colored grey.

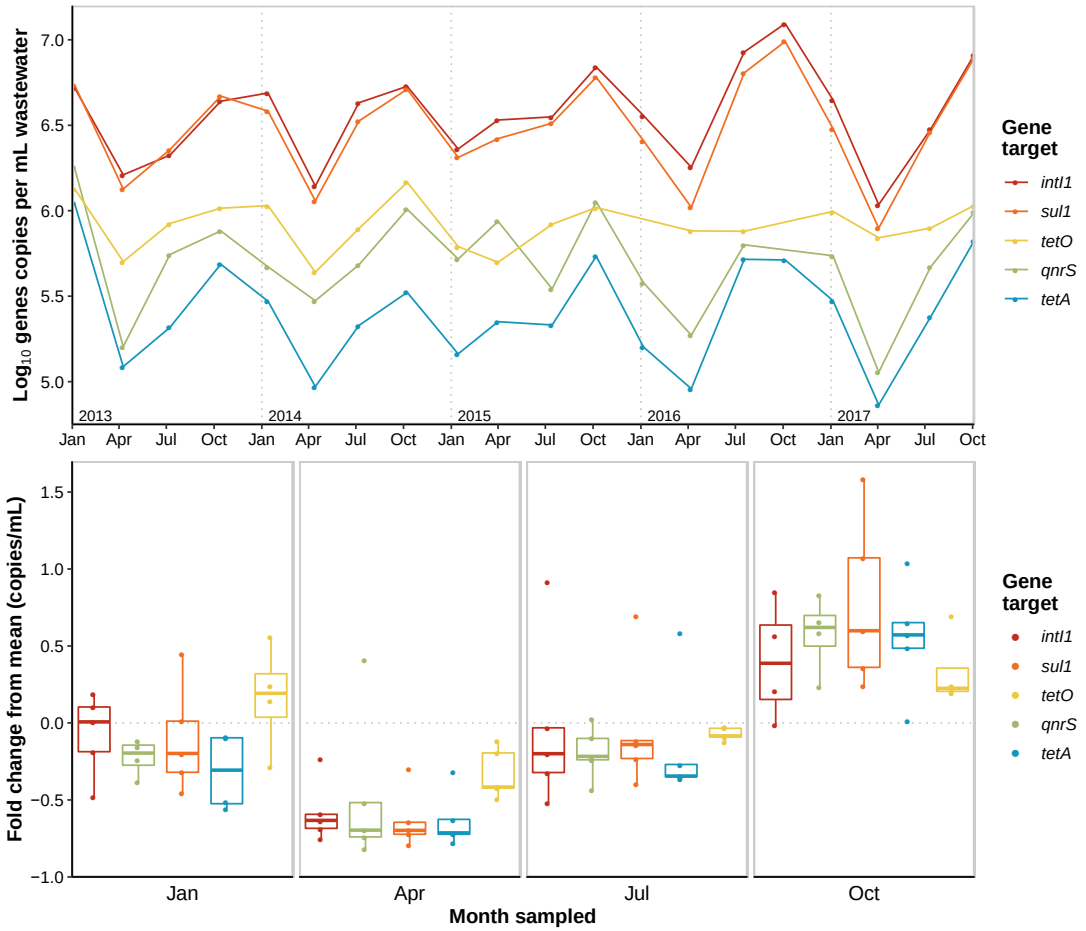

Figure 13. Resistance gene concentrations in wastewater over five years.

Concentrations (copies per mL wastewater) of ARGs (*sulI*, orange; *tetO*, yellow; *qnrS*, green; and *tet(A)*, blue) and the class 1 integron integrase, intI1 (red) measured by droplet digital PCR (ddPCR) in quarterly (Jan, Apr, Jul, Oct) wastewater samples collected from 2013 to 2017. (top) Log_{10} ARG concentrations over time. (bottom) ARG concentrations fold change from mean (dotted grey line) in each quarter. Boxes depict the median and first and third quartiles, whisker lines extend to interquartile ranges • 1.5, and points are outlier values. Fold change of 1 means that observation was 1-fold, or 2x, different from the mean of that ARG, 0 means that observation was equal to the mean.

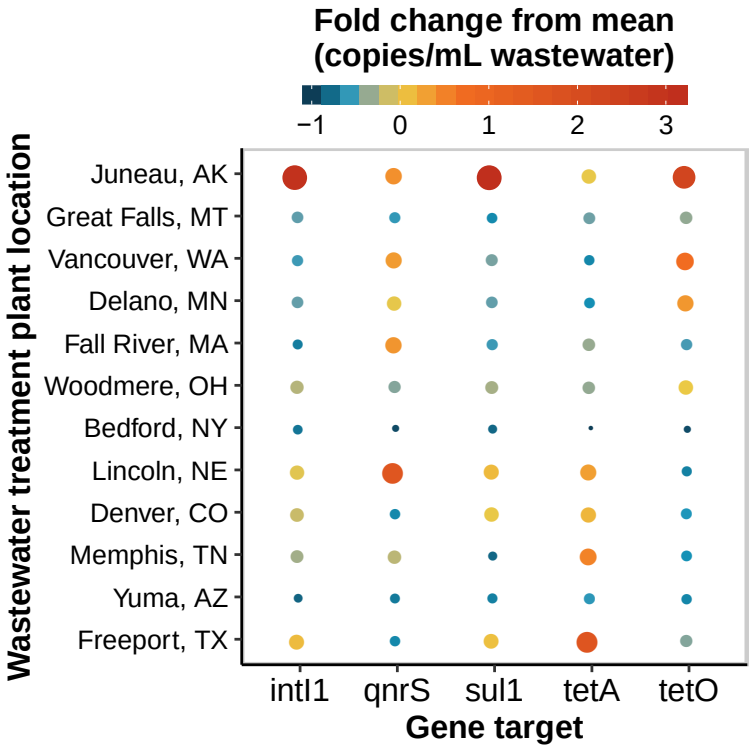

Figure 14. Resistance gene concentrations in wastewater across the USA.

ARG concentrations (copies per mL wastewater), measured by ddPCR, fold change from mean at different treatment plants across the US. Color and size of points show fold decrease (min = -1, blue) and increase (max = 3, red) from mean (0, yellow). Y-axis WWTPs ranked from top to bottom in decreasing latitude.

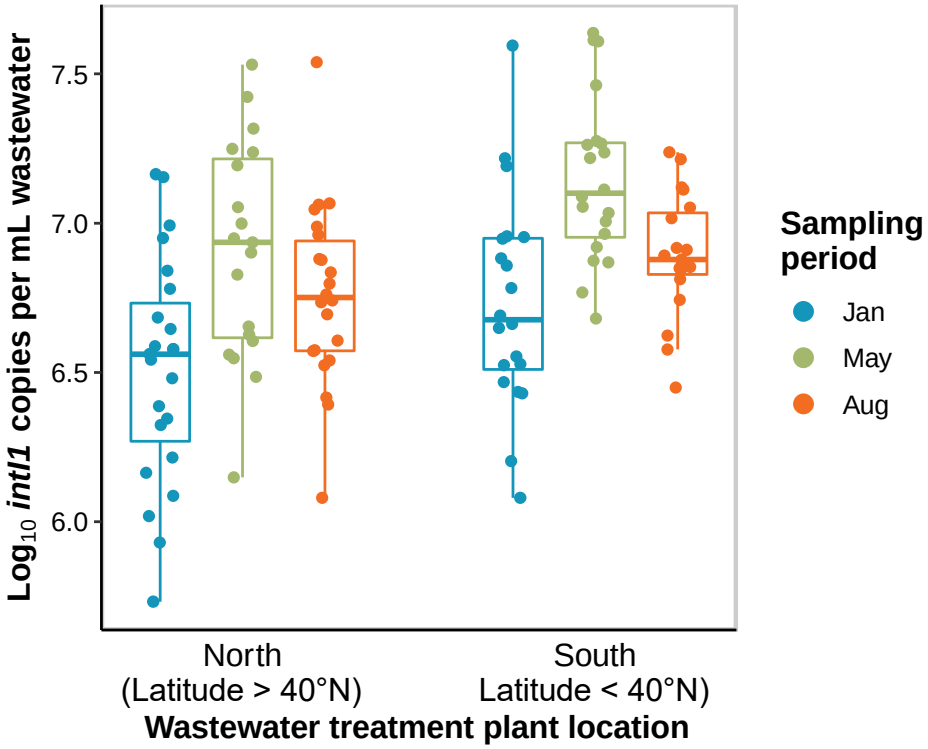

Figure 15. Seasonal concentrations of *intI1* in northern and southern USA wastewater.

Concentrations (log_{10} copies per mL wastewater) of *intI1* in treatment plants in northern US (latitude > 40°N) and in southern US (< 40°N). Each WWTP was sampled in August 2012 (orange), January 2013 (blue), and May 2013 (green). Points indicate individual samples. Boxes depict the median and first and third quartiles, whisker lines extend to interquartile ranges • 1.5.

Tables

Table 8. Metagenome sample information.

Treatment plant city, state	Date(s) sampled
JI, Milwaukee, WI	1-Aug-2012, 5-Jan-2013, 1-May-2013, 9-Apr-2018, 23-Apr-2018, 7-May-2018, 22-May-2018, 10-Sep-2018, 24-Sep-2018, 8-Oct-2018, 22-Oct-2018
SS, Milwaukee, WI	1-Aug-2012, 5-Jan-2013, 1-May-2013
Palo Alto, CA	1-Aug-2012, 5-Jan-2013, 1-May-2013
Key West, FL	1-Aug-2012, 5-Jan-2013, 1-May-2013
Sand Island, HI	10-Oct-2018, 16-Oct-2018, 14-Nov-2018, 28-Nov-2018
Gloucester, MA	1-Aug-2012, 5-Jan-2013, 1-May-2013
Laramie, WY	1-Aug-2012, 5-Jan-2013, 1-May-2013

JI = Jones Island Water Reclamation Facility
SS = South Shore Water Reclamation Facility

Table 9. ddPCR sample information.

City	State	Month(s) and year(s)	Gene(s) assayed
Milwaukee	Wisconsin	Jan, Apr, Jul, Oct; 2013, 2014, 2015, 2016, 2017	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Juneau	Alaska	Aug 2012, Jan 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Yuma	Arizona	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
North Denver	Colorado	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Fall River	Massachusetts	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Delano	Minnesota	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Great Falls	Montana	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Lincoln	Nebraska	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Bedford	New York	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Woodmere	Ohio	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Memphis	Tennessee	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Freeport	Texas	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Vancouver	Washington	Aug 2012, Jan 2013, May 2013	*intI1*, *sul1*, *qnrS*, *tetA*, *tetO*
Discovery Bay	California	Aug 2012	*intI1*
Richmond	California	Aug 2012	*intI1*
Santa Barbara 1	California	Aug 2012	*intI1*
Santa Barbara 2	California	Aug 2012	*intI1*
Stockton	California	Aug 2012	*intI1*
South Denver	Colorado	Aug 2012	*intI1*
Palm Beach	Florida	Aug 2012	*intI1*
Palmetto	Florida	Aug 2012	*intI1*
Johns Creek	Georgia	Aug 2012	*intI1*
Boonville	Indiana	Aug 2012	*intI1*
Iowa City	Iowa	Aug 2012	*intI1*
Salina	Kansas	Aug 2012	*intI1*
Hardinsburg	Kentucky	Aug 2012	*intI1*
Brockton	Massachusetts	Aug 2012	*intI1*
Gloucester	Massachusetts	Aug 2012	*intI1*
Albertville	Minnesota	Aug 2012	*intI1*

City	State	Month(s) and year(s)	Gene(s) assayed
Farmington	Minnesota	Aug 2012	*intI1*
Monticello	Minnesota	Aug 2012	*intI1*
Shakopee	Minnesota	Aug 2012	*intI1*
St. Paul	Minnesota	Aug 2012	*intI1*
Bozeman	Montana	Aug 2012	*intI1*
Hillburn	New York	Aug 2012	*intI1*
Poughkeepsie	New York	Aug 2012	*intI1*
Syracuse	New York	Aug 2012	*intI1*
Springboro	Ohio	Aug 2012	*intI1*
Burkburnett	Texas	Aug 2012	*intI1*
Kenedy	Texas	Aug 2012	*intI1*
Madison	Wisconsin	Aug 2012	*intI1*
Laramie	Wyoming	Aug 2012	*intI1*

Table 10. ddPCR gene targets.

Gene	Role	Sequence		Amplicon length	Anneal temp	Source
intI1	gene mobility	forward	CGAACGAGTGGCGGAGGGTG	311 bp	60°C	Gillings, M. R., et al. (2015). The ISME journal, 9(6), 1269-1279.
		reverse	TACCCGAGAGCTTGGCACCCA			
qnrS	quinolone resistance	forward	GACGTGCTAACTTGCGTGAT	240 bp	62°C	Rodriguez-Mozaz, S., et al. Water research, 69, 234-242.
		reverse	TGGCATTGTTGGAAACTTG			
sulI	sulfonamide resistance	forward	CGCACCGGAAACATCGCTGCAC	67 bp	60°C	Rodriguez-Mozaz, Sara, et al. Water research 69 (2015): 234-242.
		reverse	TGAAGTTCCGCCGCAAGGCTCG			
tet(A)	tetracycline resistance	forward	GCTACATCCTGCTTGCCTTC	210 bp	60°C	LaPara, T. M., et al. (2011). Environmental science & technology, 45(22), 9543-9549.
		reverse	CATAGATCGCCGTGAAGAGG			
tetO	tetracycline resistance	forward	ACGGARAGTTTATTGTATACC	171 bp	60°C	Munir, M., et al. (2011). Water research, 45(2), 681-693.
		reverse	TGGCGTATCTATAATGTTGAC			

CHAPTER IV

CONCLUSIONS

Microorganisms were the first forms of life on earth. Humans did not start significantly altering landscapes until just a couple hundred years ago, yet the effects have changed all terrestrial, atmospheric, and aquatic biomes. A new field of science, urban ecology, was developed to focus on life that has adapted to built environments. Among them, sewer systems are expansive networks of pipes that collect wastewater from municipalities, and despite getting little to no sunlight, abundant surface area, water, and nutrients have enabled organisms, particularly microorganisms, to colonize them.

It is not yet known from where resident sewer microorganisms are derived. Sources are likely some combination of soil, natural waters, stormwater, and human, industrial, and agricultural waste. A great deal of research is done in these source microbiomes, as well as within wastewater treatment plants, but little focus has been put into what is between them, the pipes. This is surprising, given that 900 billion gallons of raw sewage is leaked into the environment each year[24], carrying risks to environmental and public health. Microbes in sewers also contribute to pipe corrosion, a $1 billion a year challenge in the US[148]. Only very recently was input to wastewater treatment plants not considered a random phenomenon, showing that there are diurnal cycles to flow that changes microbial community assembly[129] and that assembly affects wastewater treatment processes[93].

Another compounding feature of sewage is its potential to facilitate the evolution and spread of microbial antibiotic resistance. Antibiotics and other stressors put selective pressures on antibiotic resistance genes (ARGs), so they are maintained in bacterial populations during conveyance. ARGs often persist or even become enriched in wastewater treatment[184–186] and are considered contaminants of emerging concern[187–189].

This dissertation used an ecological approach to characterize the urban sewer ecosystem. First, sewage microbial communities exhibited a predictable response to seasons in their assembly, a phenomenon typically attributed to surface water habitats. Second, a similar, climate-driven response was seen in ARGs that suggests proliferation with the bloom of warm-water-associated microbial communities. Broader implications to this dissertation are four-fold.

I. Microbial adaptation to built environments

Aim 1, Chapter II profiled major bacterial groups in sewage. It also individually analyzed resident-sewer and human-associated communities. This supports what was suggested previously[4,87,100] that sewage is not comprised entirely of its source microbes, and that it harbors its own groups unique to the system. To explore this further, a pangenome analysis compares genomes from individuals in a phylogenetic group, to assess core/shared genes and accessory/unique genes. This approach could reveal the closest relatives to sewer-specific taxa and the traits they maintained or evolved to adapt to this environment. A pangenome analysis of sewer *Flavobacteria* is proposed in **Appendix B**.

II. Wastewater treatment and infrastructure

Wastewater treatment is an energy intensive process that uses up to 35% of the US energy budget[190]. Aeration is typically the largest energy consumer[191]. Aeration is required to maintain activated sludge, which cultivates flocs of microorganisms that clump solids and phosphorus for easier removal. Many dominant bacterial groups in sewage can floc, including *Trichococcus*, *Acinetobacter*, and *Flavobacteria*[192]. **Aim 1, Chapter II** showed a seasonal predictability in the bacterial community assembly, therefore it is worth looking into how this affects the efficacy of activated sludge. Of note, *Trichococcus* appeared to have a negative relationship with temperature, and was completely absent from wastewater samples from southern latitudes (**Appendix A**, Fig. 16).

Wastewater collection systems are also impacted by microbial processes. Microbial-induced corrosion has accelerated in recent years due to increasing temperatures and use of sulfate-containing detergents[193]. Together, sulfur-reducing and sulfur-oxidizing bacteria convert sulfates to sulfuric acid, which disintegrates pipe material[194]. Important questions are if sewage bacteria can be used to control pipe corrosion, and how seasonality of community assembly influences decay rates.

III. Pollution tracking and risk assessment

Fluctuating microbial groups in sewage could impact how fecal pollution tracking is planned and risk assessments are made. With certain organisms more abundant at certain times of year, it is possible that targets for tracking sewage overflows could be refined specifically to seasons. Also,

if levels of high-risk groups vary throughout the year, risk assessments could be improved with this information.

IV. Dissemination of antibiotic resistance

There are many knowledge gaps in environmental antibiotic resistance, particularly in mechanisms that evolved since clinical use began. However, consensus exists in the opinion that wastewater is a critical reservoir for antibiotic resistance, whether it directly promotes evolution, serves as a sink for ARGs, or provides means of conveyance for resistant pathogens. This dissertation showed a close coupling of ARGs to environmental, temperature-dependent microbial communities. More research is needed in how these ARGs are carried (i.e., chromosomally or on mobile genetic elements) and who is carrying them (i.e., phylogeny of bacterial hosts) to scrutinize this observation any further. At the very least, it reveals an ecological dynamic to antibiotic resistance that, to the best of our knowledge, has not been seen previously. Continued longitudinal monitoring of wastewater ARGs would be helpful in catching emerging ARGs before they disseminate to waterways, animals, humans, and eventually, around the world.

Temperature dependence is clearly driving large-scale changes to bacterial communities in urban sewers. A five-year time series of Milwaukee sewage revealed that both bacteria and ARGs assembled into two distinct steady states in fall and spring. Sewage from across the US exhibited similar temperature-driven trends explained by differences in climate. Applying microbial ecological theory to sewer systems can enhance knowledge in many important topics, including wastewater surveillance and epidemiology, infrastructure maintenance, water treatment, and

antibiotic resistance. Learning from sewer microorganisms about colonization, resilience, and adaptation is imperative in a progressively stressed and urbanized world.

REFERENCES

1. Johnson, S. *The ghost map: The story of London's most terrifying epidemic--and how it changed science, cities, and the modern world*. (Penguin, 2006).
2. Gaynes, R. P. *Germ theory: medical pioneers in infectious diseases*. (John Wiley & Sons, 2011).
3. Escolà Casas, M. *et al.* Showcasing the potential of wastewater-based epidemiology to track pharmaceuticals consumption in cities: Comparison against prescription data collected at fine spatial resolution. *Environ. Int.* **150**, 106404 (2021).
4. Newton, R. J. *et al.* Sewage reflects the microbiomes of human populations. *MBio* **6**, 1–9 (2015).
5. Sulej-Suchomska, A. M. *et al.* Urban wastewater analysis as an effective tool for monitoring illegal drugs, including new psychoactive substances, in the Eastern European region. *Sci. Rep.* **10**, 4885 (2020).
6. Bogler, A. *et al.* Rethinking wastewater risks and monitoring in light of the COVID-19 pandemic. *Nat. Sustain.* **3**, 981–990 (2020).
7. Hoar, C. *et al.* Looking Forward: The Role of Academic Researchers in Building Sustainable Wastewater Surveillance Programs. (2022).
8. Hillary, L. S., Malham, S. K., McDonald, J. E. & Jones, D. L. Wastewater and public health: the potential of wastewater surveillance for monitoring COVID-19. *Curr. Opin. Environ. Sci. Heal.* **17**, 14–20 (2020).
9. Daughton, C. G. Wastewater surveillance for population-wide Covid-19: the present and future. *Sci. Total Environ.* **736**, 139631 (2020).

10. Randazzo, W., Cuevas-Ferrando, E., Sanjuán, R., Domingo-Calap, P. & Sánchez, G. Metropolitan wastewater analysis for COVID-19 epidemiological surveillance. *Int. J. Hyg. Environ. Health* **230**, 113621 (2020).

11. Xiao, A. *et al.* Metrics to relate COVID-19 wastewater data to clinical testing dynamics. *Water Res.* **212**, 118070 (2022).

12. Naughton, C. C. *et al.* Show us the data: global COVID-19 wastewater monitoring efforts, equity, and gaps. *MedRXiv* (2021).

13. McMahan, C. S. *et al.* COVID-19 wastewater epidemiology: a model to estimate infected populations. *Lancet Planet. Heal.* **5**, e874–e881 (2021).

14. Keshaviah, A. *et al.* Separating Signal from Noise in Wastewater Data: An Algorithm to Identify Community-Level COVID-19 Surges. *medRxiv* (2022).

15. Gormley, M., Aspray, T. J. & Kelly, D. A. COVID-19: mitigating transmission via wastewater plumbing systems. *Lancet Glob. Heal.* **8**, e643 (2020).

16. Schussman, M. K. & McLellan, S. L. Effect of time and temperature on SARS-COV-2 in municipal wastewater conveyance systems. *Water* **14**, 1373 (2022).

17. Nghiem, L. D., Morgan, B., Donner, E. & Short, M. D. The COVID-19 pandemic: considerations for the waste and wastewater services sector. *Case Stud. Chem. Environ. Eng.* **1**, 100006 (2020).

18. Lahrich, S. *et al.* Review on the contamination of wastewater by COVID-19 virus: Impact and treatment. *Sci. Total Environ.* **751**, 142325 (2021).

19. Goldstein, J. & Otterman, S. Polio Has Been Detected in New York City Wastewater, Officials Say. *New York Times* (2022).

20. Hopkins, J. N. N. The Cloaca Maxima and the monumental manipulation of water in archaic Rome. *waters Rome* **4**, 1–15 (2007).

21. Burian, S. J., Nix, S. J., Pitt, R. E. & Durrans, S. R. Urban wastewater management in the United States: Past, present, and future. *J. Urban Technol.* **7**, 33–62 (2000).

22. Engineers, A. S. of C. Infrastructure report card. in (American Society of Civil Engineers Reston, VA, 2017).

23. US Environmental Protection Agency. The Sources and Solutions: Wastewater. *Nutrient Pollution* (2022). Available at: https://www.epa.gov/nutrientpollution/sources-and-solutions-wastewater. (Accessed: 11th January 2022)

24. Engineers, A. S. of C. Failure to Act: The Economic Impact of Current Investment Trends in Water and Wastewater Treatment Infrastructure. in (American Society of Civil Engineers, 2011).

25. CDC. Centers for disease control and prevention, national center for emerging and zoonotic infectious diseases (NCEZID), Division of Foodborne, Waterborne, and Environmental Diseases (DFWED). (2012).

26. Mac Kenzie, W. R. *et al.* A massive outbreak in Milwaukee of Cryptosporidium infection transmitted through the public water supply. *N. Engl. J. Med.* **331**, 161–167 (1994).

27. SINGH, R. Water Management and Conservation Practices in Indus Valley Civilization.

28. Angelakis, A. N., Asano, T., Bahri, A., Jimenez, B. E. & Tchobanoglous, G. Water reuse: from ancient to modern times and the future. *Front. Environ. Sci.* 26 (2018).

29. Symons, G. E. Water treatment through the ages. *Journal-American Water Work. Assoc.* **98**, 87–98 (2006).

30. Cancer, I. A. for R. on. Chlorinated drinking-water; chlorination by-products; some other halogenated compounds; cobalt and cobalt compounds. *IARC Monogr. Eval. Carcinog. risks to humans* **52**, (1991).

31. Ardern, E. & Lockett, W. T. Experiments on the oxidation of sewage without the aid of filters. *J. Soc. Chem. Ind.* **33**, 523–539 (1914).

32. US Environmental Protection Agency. Basic Information about Anaerobic Digestion (AD). *Anaerobic Digestion* Available at: https://www.epa.gov/anaerobic-digestion/basic-information-about-anaerobic-digestion-ad. (Accessed: 11th August 2022)

33. Maktabifard, M., Zaborowska, E. & Makinia, J. Achieving energy neutrality in wastewater treatment plants through energy savings and enhancing renewable energy production. *Rev. Environ. Sci. Bio/Technology* **17**, 655–689 (2018).

34. Ho, S. W., Cheung, K. K. & Fung, W. C. Sustainable wastewater treatment – ways to achieve energy neutrality. *HKIE Trans.* **21**, 240–252 (2014).

35. ASME Milwaukee. Using Digester Gas as Fuel. *History & Heritage* (2007). Available at: http://asmemilwaukee.org/38-digester.html.

36. US Environmental Protection Agency. How Wastewater Treatment Works... The Basics. (1998).

37. NOAA. How Much Oxygen Comes from the Ocean? (2020).

38. Malin, G. Sulphur, climate and the microbial maze. *Nature* **387**, 857–858 (1997).

39. Khoruts, A. & Sadowsky, M. J. Understanding the mechanisms of faecal microbiota transplantation. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 508–516 (2016).

40. Walters, W. A., Xu, Z. & Knight, R. Meta-analyses of human gut microbes associated with

obesity and IBD. *FEBS Lett.* **588**, 4223–4233 (2014).

41. Valles-Colomer, M. *et al.* The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat. Microbiol.* **4**, 623–632 (2019).

42. Sender, R., Fuchs, S. & Milo, R. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol.* **14**, e1002533 (2016).

43. Steel, E. A., Kennedy, M. C., Cunningham, P. G. & Stanovick, J. S. Applied statistics in ecology: common pitfalls and simple solutions. *Ecosphere* **4**, 1–13 (2013).

44. McLellan, S. L., Daniels, A. D. & Salmore, A. K. Genetic characterization of Escherichia coli populations from host sources of fecal pollution by using DNA fingerprinting. *Appl. Environ. Microbiol.* **69**, 2587–2594 (2003).

45. Templar, H. A., Dila, D. K., Bootsma, M. J., Corsi, S. R. & McLellan, S. L. Quantification of human-associated fecal indicators reveal sewage from urban watersheds as a source of pollution to Lake Michigan. *Water Res.* **100**, 556–567 (2016).

46. Bower, P. A., Scopel, C. O., Jensen, E. T., Depas, M. M. & McLellan, S. L. Detection of genetic markers of fecal indicator bacteria in Lake Michigan and determination of their relationship to Escherichia coli densities using standard microbiological methods. *Appl. Environ. Microbiol.* **71**, 8305–8313 (2005).

47. Heather, J. M. & Chain, B. The sequence of sequencers: The history of sequencing DNA. *Genomics* **107**, 1–8 (2016).

48. Gusella, J. F. *et al.* A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* **306**, 234–238 (1983).

49. Brock, D. C. & Moore, G. E. *Understanding Moore's law: four decades of innovation*.

(Chemical Heritage Foundation, 2006).

50. Muir, P. *et al.* The real cost of sequencing: scaling computation to keep pace with data generation. *Genome Biol.* **17**, 1–9 (2016).

51. Behjati, S. & Tarpey, P. S. What is next generation sequencing? *Arch. Dis. Childhood-Education Pract.* **98**, 236–238 (2013).

52. edu, C. S. and A. C. W. R. H. waterston@ gs. washington. edu L. E. S. lander@ broad. mit. edu W. R. K. rwilson@ watson. wustl. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87 (2005).

53. Woese, C. R. & Fox, G. E. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci.* **74**, 5088–5090 (1977).

54. Pace, N. R., Sapp, J. & Goldenfeld, N. Phylogeny and beyond: Scientific, historical, and conceptual significance of the first tree of life. *Proc. Natl. Acad. Sci.* **109**, 1011–1018 (2012).

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

56. Knight, R. *et al.* Best practices for analysing microbiomes. *Nat. Rev. Microbiol.* **16**, 410–422 (2018).

57. Pollock, J., Glendinning, L., Wisedchanwet, T. & Watson, M. The madness of microbiome: attempting to find consensus "best practice" for 16S microbiome studies. *Appl. Environ. Microbiol.* **84**, e02627-17 (2018).

58. Johnson, J. S. *et al.* Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat. Commun.* **10**, 1–11 (2019).

59. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* **108**, 4516–4522 (2011).

60. Costello, E. K. *et al.* Bacterial community variation in human body habitats across space and time. *Science (80-.).* **326**, 1694–1697 (2009).

61. Roesch, L. F. W. *et al.* Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J.* **1**, 283–290 (2007).

62. Sogin, M. L. *et al.* Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci.* **103**, 12115–12120 (2006).

63. Ley, R. E., Lozupone, C. A., Hamady, M., Knight, R. & Gordon, J. I. Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat. Rev. Microbiol.* **6**, 776–788 (2008).

64. Hutchings, M. I., Truman, A. W. & Wilkinson, B. Antibiotics: past, present and future. *Curr. Opin. Microbiol.* **51**, 72–80 (2019).

65. Prestinaci, F., Pezzotti, P. & Pantosti, A. Antimicrobial resistance: a global multifaceted phenomenon. *Pathog. Glob. Health* **109**, 309–318 (2015).

66. Organization, W. H. global health issues to track in 2021. *WHO. Newsroom* **24**, 2020 (10AD).

67. Sabol, K. *et al.* Emergence of daptomycin resistance in Enterococcus faecium during daptomycin therapy. *Antimicrob. Agents Chemother.* **49**, 1664–1665 (2005).

68. Seifert, R. & Schirmer, B. Problems associated with the use of the term "antibiotics". *Naunyn. Schmiedebergs. Arch. Pharmacol.* **394**, 2153–2166 (2021).

69. Allen, H. K. *et al.* Call of the wild: antibiotic resistance genes in natural environments. *Nat. Rev. Microbiol.* **8**, 251–259 (2010).

70. Davies, J. Are antibiotics naturally antibiotics? *J. Ind. Microbiol. Biotechnol.* **33**, 496–499 (2006).

71. Sengupta, S., Chattopadhyay, M. K. & Grossart, H.-P. The multifaceted roles of antibiotics and antibiotic resistance in nature. *Front. Microbiol.* **4**, 47 (2013).

72. Andersson, D. I. & Hughes, D. Microbiological effects of sublethal levels of antibiotics. **12**, 465–479 (2014).

73. Linares, J. F., Gustafsson, I., Baquero, F. & Martinez, J. L. Antibiotics as intermicrobial signaling agents instead of weapons. *Proc. Natl. Acad. Sci.* **103**, 19484–19489 (2006).

74. Davies, J. & Davies, D. Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* **74**, 417–433 (2010).

75. Goh, E.-B. *et al.* Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proc. Natl. Acad. Sci.* **99**, 17025–17030 (2002).

76. Gutierrez, A. *et al.* β-Lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity. *Nat. Commun.* **4**, 1–9 (2013).

77. Beaber, J. W., Hochhut, B. & Waldor, M. K. SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* **427**, 72–74 (2004).

78. Miller, C. *et al.* SOS response induction by ß-lactams and bacterial defense against antibiotic lethality. *Science (80-.).* **305**, 1629–1631 (2004).

79. Maiques, E. *et al.* β-Lactam antibiotics induce the SOS response and horizontal transfer of virulence factors in Staphylococcus aureus. *J. Bacteriol.* **188**, 2726–2729 (2006).

80. Dionisio, F., Matic, I., Radman, M., Rodrigues, O. R. & Taddei, F. Plasmids spread very fast in heterogeneous bacterial communities. *Genetics* **162**, 1525–1532 (2002).

81. Partridge, S. R., Tsafnat, G., Coiera, E. & Iredell, J. R. Gene cassettes and cassette arrays in mobile resistance integrons. *FEMS Microbiol. Rev.* **33**, 757–784 (2009).

82. Gilbert, P., Maira-Litran, T., McBain, A. J., Rickard, A. H. & Whyte, F. W. The physiology and collective recalcitrance of microbial biofilm communities. (2002).

83. Gullberg, E., Albrecht, L. M., Karlsson, C., Sandegren, L. & Andersson, D. I. Selection of a multidrug resistance plasmid by sublethal levels of antibiotics and heavy metals. *MBio* **5**, e01918-14 (2014).

84. Lundström, S. V *et al.* Minimal selective concentrations of tetracycline in complex aquatic bacterial biofilms. *Sci. Total Environ.* **553**, 587–595 (2016).

85. Martínez, J. L., Coque, T. M. & Baquero, F. What is a resistance gene? Ranking risk in resistomes. *Nat. Rev. Microbiol.* **13**, 116–123 (2015).

86. Bengtsson-palme, J., Kristiansson, E. & Larsson, D. G. J. Environmental factors influencing the development and spread of antibiotic resistance. 68–80 (2018). doi:10.1093/femsre/fux053

87. Fisher, J. C., Levican, A., Figueras, M. J. & McLellan, S. L. Population dynamics and ecology of Arcobacter in sewage. *Front. Microbiol.* **5**, 525 (2014).

88. Skwor, T. *et al.* Prevalence of Potentially Pathogenic Antibiotic-Resistant Aeromonas spp. in Treated Urban Wastewater Effluents versus Recipient Riverine Populations: a 3-Year Comparative Study. *Appl. Environ. Microbiol.* **86**, (2020).

89. Da Costa, P. M., Vaz-Pires, P. & Bernardo, F. Antimicrobial resistance in Enterococcus spp. isolated in inflow, effluent and sludge from municipal sewage water treatment plants. *Water Res.* **40**, 1735–1740 (2006).

90. Ahmed, W. *et al.* Antibiotic resistance and sewage-associated marker genes in untreated sewage and a river characterized during baseflow and stormflow. *Front. Microbiol.* **12**, 632850 (2021).

91. Pruden, A. Balancing water sustainability and public health goals in the face of growing concerns about antibiotic resistance. (2014).

92. Rhodes, G. *et al.* Distribution of oxytetracycline resistance plasmids between aeromonads in hospital and aquaculture environments: implication of Tn 1721 in dissemination of the tetracycline resistance determinant Tet A. *Appl. Environ. Microbiol.* **66**, 3883–3890 (2000).

93. Dottorini, G. *et al.* Mass-immigration determines the assembly of activated sludge microbial communities. *Proc. Natl. Acad. Sci.* **118**, e2021589118 (2021).

94. Wade, T. J. *Report on 2009 national epidemiologic and environmental assessment of recreational water epidemiology studies*. (United States Environmental Protection Agency, Office of Research and …, 2010).

95. Cayford, B. I., Dennis, P. G., Keller, J., Tyson, G. W. & Bond, P. L. High-throughput amplicon sequencing reveals distinct communities within a corroding concrete sewer system. *Appl. Environ. Microbiol.* **78**, 7160–7162 (2012).

96. Ling, A. L. *et al.* Carbon dioxide and hydrogen sulfide associations with regional bacterial diversity patterns in microbially induced concrete corrosion. *Environ. Sci. Technol.* **48**, 7357–7364 (2014).

97. Ling, A. L. *et al.* High-resolution microbial community succession of microbially induced concrete corrosion in working sanitary manholes. *PLoS One* **10**, e0116400 (2015).

98. Kip, N. & Van Veen, J. A. The dual role of microbes in corrosion. *ISME J.* **9**, 542–551

(2015).

99. Li, X., Kappler, U., Jiang, G. & Bond, P. L. The ecology of acidophilic microorganisms in the corroding concrete sewer environment. *Front. Microbiol.* **8**, 1–16 (2017).

100. McLellan, S. L., Huse, S. M., Mueller-Spitz, S. R., Andreishcheva, E. N. & Sogin, M. Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. *Environ. Microbiol.* **12**, 378–392 (2010).

101. Roguet, A., Newton, R. J., Eren, A. M. & McLellan, S. L. Guts of the Urban Ecosystem: Microbial Ecology of Sewer Infrastructure. *Msystems* **7**, e00118-22 (2022).

102. Wagner, M. *et al.* Microbial community composition and function in wastewater treatment plants. *Antonie Van Leeuwenhoek* **81**, 665–680 (2002).

103. Zhang, L., Shen, Z., Fang, W. & Gao, G. Composition of bacterial communities in municipal wastewater treatment plant. *Sci. Total Environ.* **689**, 1181–1191 (2019).

104. Wu, L. *et al.* Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nat. Microbiol.* **4**, 1183–1195 (2019).

105. Cydzik-Kwiatkowska, A. & Zielińska, M. Bacterial communities in full-scale wastewater treatment systems. *World J. Microbiol. Biotechnol.* **32**, 1–8 (2016).

106. Al Ali, A. A., Naddeo, V., Hasan, S. W. & Yousef, A. F. Correlation between bacterial community structure and performance efficiency of a full-scale wastewater treatment plant. *J. Water Process Eng.* **37**, 101472 (2020).

107. Daims, H., Taylor, M. W. & Wagner, M. Wastewater treatment: a model system for microbial ecology. *Trends Biotechnol.* **24**, 483–489 (2006).

108. Jiang, X.-T., Guo, F. & Zhang, T. Population dynamics of bulking and foaming bacteria in

a full-scale wastewater treatment plant over five years. *Sci. Rep.* **6**, 1–9 (2016).

109. Lee, S. H., Kang, H. J. & Park, H. D. Influence of influent wastewater communities on temporal variation of activated sludge communities. *Water Res.* **73**, 132–144 (2015).

110. Wells, G. F. *et al.* Microbial biogeography across a full-scale wastewater treatment plant transect: Evidence for immigration between coupled processes. *Appl. Microbiol. Biotechnol.* **98**, 4723–4736 (2014).

111. Gray, N. D., Miskin, I. P., Kornilova, O., Curtis, T. P. & Head, I. M. Occurrence and activity of archaea in aerated activated sludge wastewater treatment plants. *Environ. Microbiol.* **4**, 158–168 (2002).

112. McLellan, S. L. & Roguet, A. The unexpected habitat in sewer pipes for the propagation of microbial communities and their imprint on urban waters. *Curr. Opin. Biotechnol.* **57**, 34–41 (2019).

113. Millar, J. A. & Raghavan, R. Accumulation and expression of multiple antibiotic resistance genes in *Arcobacter cryaerophilus* that thrives in sewage. *PeerJ* (2017). doi:10.7717/peerj.3269

114. Johnning, A. *et al.* Acquired genetic mechanisms of a multiresistant bacterium isolated from a treatment plant receiving wastewater from antibiotic production. *Appl. Environ. Microbiol.* **79**, 7256–7263 (2013).

115. Marathe, N. P., Shetty, S. A., Shouche, Y. S. & Larsson, D. G. J. Limited bacterial diversity within a treatment plant receiving antibiotic containing waste from bulk drug production. *PLoS One* **11**, 1–12 (2016).

116. Marti, E., Jofre, J. & Balcazar, J. L. Prevalence of Antibiotic Resistance Genes and Bacterial

Community Composition in a River Influenced by a Wastewater Treatment Plant. *PLoS One* (2013). doi:10.1371/journal.pone.0078906

117. Munir, M., Wong, K. & Xagoraraki, I. Release of antibiotic resistant bacteria and genes in the effluent and biosolids of five wastewater utilities in Michigan. *Water Res.* **45**, 681–693 (2011).

118. Lapara, T. M. *et al.* Tertiary-treated municipal wastewater is a significant point source of antibiotic resistance genes into Duluth-Superior Harbor. *Environ. Sci. Technol.* (2011). doi:10.1021/es202775r

119. Barrett, M. H. & Reynolds, J. H. a Review of the Effects of Sewer Leakage on G R 0 U N D Wate R Quality. *Journal* **17 N1**, (2003).

120. Ly, D. K. & Chui, T. F. M. Modeling sewage leakage to surrounding groundwater and stormwater drains. *Water Sci. Technol.* **66**, 2659–2665 (2012).

121. Roehrdanz, P. R. *et al.* Spatial Models of Sewer Pipe Leakage Predict the Occurrence of Wastewater Indicators in Shallow Urban Groundwater. *Environ. Sci. Technol.* **51**, 1213–1223 (2017).

122. Jalliffier-Verne, I. *et al.* Cumulative effects of fecal contamination from combined sewer overflows: Management for source water protection. *J. Environ. Manage.* **174**, 62–70 (2016).

123. Huang, C. *et al.* Environmental transport of emerging human-pathogenic Cryptosporidium species and subtypes through combined sewer overflow and wastewater. *Appl. Environ. Microbiol.* **83**, e00682-17 (2017).

124. DeFlorio-Barker, S., Wing, C., Jones, R. M. & Dorevitch, S. Estimate of incidence and cost

of recreational waterborne illness on United States surface waters. *Environ. Heal.* **17**, 3 (2018).

125. Linz, A. M. *et al.* Bacterial community composition and dynamics spanning five years in freshwater bog lakes. *MSphere* **2**, e00169-17 (2017).

126. Fuhrman, J. A. *et al.* Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 13104–13109 (2006).

127. Gilbert, J. A. *et al.* Defining seasonal marine microbial community dynamics. *ISME J.* **6**, 298–308 (2012).

128. Kara, E. L., Hanson, P. C., Hu, Y. H., Winslow, L. & McMahon, K. D. A decade of seasonal dynamics and co-occurrences within freshwater bacterioplankton communities from eutrophic Lake Mendota, WI, USA. *ISME J.* **7**, 680–684 (2013).

129. Guo, B., Liu, C., Gibson, C. & Frigon, D. Wastewater microbial community structure and functional traits change over short timescales. *Sci. Total Environ.* **662**, 779–785 (2019).

130. Huse, S. M. *et al.* Comparison of brush and biopsy sampling methods of the ileal pouch for assessment of mucosa-associated microbiota of human subjects. *Microbiome* **2**, 5 (2014).

131. Andrews, S. FastQC: a quality control tool for high throughput sequence data. (2010).

132. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).

133. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

134. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, 590–596 (2013).

135. Davis, N. M., Proctor, Di. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 1–14 (2018).

136. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).

137. Untergasser, A. *et al.* Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.* **35**, (2007).

138. Cole, J. R. *et al.* Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42**, (2014).

139. Green, H. C. *et al.* Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. *Appl. Environ. Microbiol.* **80**, 3086–3094 (2014).

140. Oksanen, J. *et al.* The vegan package. *Community Ecol. Packag.* **10**, 631–637 (2007).

141. Team, R. C. R: A language and environment for statistical computing. (2013).

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

143. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

144. Wickham, H. *ggplot2 Elegant Graphics for Data Analysis*. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* (2016). doi:10.1007/978-3-319-24277-4

145. Newton, R. J., Bootsma, M. J., Morrison, H. G., Sogin, M. L. & McLellan, S. L. A microbial signature approach to identify fecal pollution in the waters off an urbanized coast of Lake

Michigan. *Microb. Ecol.* **65**, 1011–1023 (2013).

146. Roguet, A., Eren, A. M., Newton, R. J. & McLellan, S. L. Fecal source identification using random forest. *Microbiome* **6**, 185 (2018).

147. Li, W., Zheng, T., Ma, Y. & Liu, J. Current status and future prospects of sewer biofilms: Their structure, influencing factors, and substance transformations. *Sci. Total Environ.* 133815 (2019).

148. Rauch, W. & Kleidorfer, M. Replace contamination, not the pipes. *Science (80-.).* **345**, 734–735 (2014).

149. Crump, B. C. *et al.* Circumpolar synchrony in big river bacterioplankton. *Proc. Natl. Acad. Sci.* **106**, 21208–21212 (2009).

150. Giovannoni, S. J. & Vergin, K. L. Seasonality in ocean microbial communities. *Science (80-.).* **335**, 671–676 (2012).

151. Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D. & Bertilsson, S. A guide to the natural history of freshwater lake bacteria. *Microbiol. Mol. Biol. Rev.* **75**, 14–49 (2011).

152. Huttenhower, C. *et al.* Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).

153. Bryskier, A. *Antimicrobial agents: antibacterials and antifungals.* (ASM press, 2005).

154. Böckelmann, U. *et al.* Quantitative PCR monitoring of antibiotic resistance genes and bacterial pathogens in three European artificial groundwater recharge systems. *Appl. Environ. Microbiol.* **75**, 154–163 (2009).

155. Börjesson, S. *et al.* Quantification of genes encoding resistance to aminoglycosides, β-lactams and tetracyclines in wastewater environments by real-time PCR. *Int. J. Environ.*

Health Res. **19**, 219–230 (2009).

156. Caucci, S. *et al.* Seasonality of antibiotic prescriptions for outpatients and resistance genes in sewers and wastewater treatment plant outflow. *FEMS Microbiol. Ecol.* **92**, fiw060 (2016).

157. 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).

158. Rodriguez-Mozaz, S. *et al.* Occurrence of antibiotics and antibiotic resistance genes in hospital and urban wastewaters and their impact on the receiving river. *Water Res.* **69**, 234–242 (2015).

159. LaMartina, E. Lou, Mohaimani, A. A. & Newton, R. J. Urban wastewater bacterial communities assemble into seasonal steady states. *Microbiome* **9**, 1–13 (2021).

160. Minoche, A. E., Dohm, J. C. & Himmelbauer, H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biol.* **12**, 1–15 (2011).

161. Eren, A. M., Vineis, J. H., Morrison, H. G. & Sogin, M. L. A filtering method to generate high quality short reads using Illumina paired-end technology. *PLoS One* **8**, e66643 (2013).

162. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).

163. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* **11**, e0163962 (2016).

164. Alcock, B. P. *et al.* CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48**, D517–D525 (2020).

165. LaPara, T. M. *et al.* Tertiary-treated municipal wastewater is a significant point source of antibiotic resistance genes into Duluth-Superior Harbor. *Environ. Sci. Technol.* **45**, 9543–9549 (2011).

166. Munir, M., Wong, K. & Xagoraraki, I. Release of antibiotic resistant bacteria and genes in the effluent and biosolids of five wastewater utilities in Michigan. *Water Res.* **45**, 681–693 (2011).

167. Gillings, M. R. *et al.* Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME J.* **9**, 1269–1279 (2015).

168. Oksanen, A. J. *et al.* The vegan Package; Community Ecology Package (Version 1.15-1). (2008).

169. De Vries, A., Ripley, B. D. & de Vries, M. A. Package 'ggdendro'. *Creat. dendrograms tree diagrams using "ggplot2* (2022).

170. Meschiari, M. *et al.* Long-Term Impact of the COVID-19 Pandemic on In-Hospital Antibiotic Consumption and Antibiotic Resistance: A Time Series Analysis (2015–2021). *Antibiotics* **11**, 826 (2022).

171. Chandy, S. J. *et al.* The impact of policy guidelines on hospital antibiotic use over a decade: a segmented time series analysis. *PLoS One* **9**, e92206 (2014).

172. Jimenez, F. *et al.* Feature selection based multivariate time series forecasting: An application to antibiotic resistance outbreaks prediction. *Artif. Intell. Med.* **104**, 101818 (2020).

173. López-Lozano, J.-M. *et al.* A nonlinear time-series analysis approach to identify thresholds in associations between population antibiotic use and rates of resistance. *Nat. Microbiol.* **4**, 1160–1172 (2019).

174. Tóth, H. *et al.* Evolution of the Gram-Negative Antibiotic Resistance Spiral over Time: A Time-Series Analysis. *Antibiotics* **10**, 734 (2021).

175. Jeffrey, B., Aanensen, D. M., Croucher, N. J. & Bhatt, S. Predicting the future distribution of antibiotic resistance using time series forecasting and geospatial modelling. *Wellcome Open Res.* **5**, 194 (2020).

176. Novo, A., André, S., Viana, P., Nunes, O. C. & Manaia, C. M. Antibiotic resistance, antimicrobial residues and bacterial community composition in urban wastewater. *Water Res.* **47**, 1875–1887 (2013).

177. Börjesson, S., Melin, S., Matussek, A. & Lindgren, P.-E. A seasonal study of the mecA gene and Staphylococcus aureus including methicillin-resistant S. aureus in a municipal wastewater treatment plant. *Water Res.* **43**, 925–932 (2009).

178. Amos, G. C. A. *et al.* Validated predictive modelling of the environmental resistome. *ISME J.* **9**, 1467–1476 (2015).

179. Wang, J., Chu, L., Wojnárovits, L. & Takács, E. Occurrence and fate of antibiotics, antibiotic resistant genes (ARGs) and antibiotic resistant bacteria (ARB) in municipal wastewater treatment plant: An overview. *Sci. Total Environ.* **744**, 140997 (2020).

180. Schloss, P. D. & Handelsman, J. Status of the microbial census. *Microbiol. Mol. Biol. Rev.* **68**, 686–691 (2004).

181. Stewart, E. J. Growing unculturable bacteria. *J. Bacteriol.* **194**, 4151–4160 (2012).

182. Zhang, Z. *et al.* Assessment of global health risk of antibiotic resistance genes. *Nat. Commun.* **13**, 1–11 (2022).

183. Walsh, T. R., Weeks, J., Livermore, D. M. & Toleman, M. A. Dissemination of NDM-1 positive bacteria in the New Delhi environment and its implications for human health: an environmental point prevalence study. *Lancet Infect. Dis.* **11**, 355–362 (2011).

184. Mao, D. *et al.* Prevalence and proliferation of antibiotic resistance genes in two municipal wastewater treatment plants. *Water Res.* **85**, 458–466 (2015).

185. Ju, F. *et al.* Antibiotic resistance genes and human bacterial pathogens: co-occurrence, removal, and enrichment in municipal sewage sludge digesters. *Water Res.* **91**, 1–10 (2016).

186. Karkman, A. *et al.* High-throughput quantification of antibiotic resistance genes from an urban wastewater treatment plant. *FEMS Microbiol. Ecol.* **92**, fiw014 (2016).

187. Berendonk, T. U. *et al.* Tackling antibiotic resistance: the environmental framework. *Nat. Rev. Microbiol.* **13**, 310–317 (2015).

188. Anand, U. *et al.* Potential environmental and human health risks caused by antibiotic-resistant bacteria (ARB), antibiotic resistance genes (ARGs) and emerging contaminants (ECs) from municipal solid waste (MSW) landfill. *Antibiotics* **10**, 374 (2021).

189. Sanderson, H. *et al.* Antimicrobial resistant genes and organisms as environmental contaminants of emerging concern: addressing global public health risks. in *Management of emerging public health issues and risks* 147–187 (Elsevier, 2019).

190. NYSERDA. Statewide Assessment of Energy Use by the Municipal Water and Wastewater Sector. *New York State Energy Research and Development Authority* (2008).

191. Rosso, D., Larson, L. E. & Stenstrom, M. K. Aeration of large-scale municipal wastewater

treatment plants: state of the art. *Water Sci. Technol.* **57**, 973–978 (2008).

192. Nielsen, P. H., Kragelund, C., Seviour, R. J. & Nielsen, J. L. Identity and ecophysiology of filamentous bacteria in activated sludge. *FEMS Microbiol. Rev.* **33**, 969–998 (2009).

193. Wells, T., Melchers, R. E. & Bond, P. Factors involved in the long term corrosion of concrete sewers. *Australas. Corros. Assoc. Proc. Corros. Prev. Coffs Harbour, Aust.* **11**, (2009).

194. Wu, M., Wang, T., Wu, K. & Kan, L. Microbiologically induced corrosion of concrete in sewer structures: A review of the mechanisms and phenomena. *Constr. Build. Mater.* **239**, 117813 (2020).

195. Sims, N. & Kasprzyk-Hordern, B. Future perspectives of wastewater-based epidemiology: monitoring infectious disease spread and resistance to the community level. *Environ. Int.* **139**, 105689 (2020).

196. Bivins, A. *et al.* Wastewater-based epidemiology: global collaborative to maximize contributions in the fight against COVID-19. (2020).

197. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

198. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* **46**, W537–W544 (2018).

199. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).

200. Schloss, P. D. *et al.* Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ.*

Microbiol. **75**, 7537–7541 (2009).

201. McBride, M. J. & Nakane, D. Flavobacterium gliding motility and the type IX secretion system. *Curr. Opin. Microbiol.* **28**, 72–77 (2015).

202. Loch, T. P. & Faisal, M. Emerging flavobacterial infections in fish: a review. *J. Adv. Res.* **6**, 283–300 (2015).

203. Zhu, Y. *et al.* Genetic analyses unravel the crucial role of a horizontally acquired alginate lyase for brown algal biomass degradation by Zobellia galactanivorans. **19**, 2164–2181 (2017).

204. Kirchman, D. L. The ecology of Cytophaga–Flavobacteria in aquatic environments. *FEMS Microbiol. Ecol.* **39**, 91–100 (2002).

205. Fandino, L. B., Riemann, L., Steward, G. F. & Azam, F. Population dynamics of Cytophaga-Flavobacteria during marine phytoplankton blooms analyzed by real-time quantitative PCR. *Aquat. Microb. Ecol.* **40**, 251–257 (2005).

206. Eiler, A. & Bertilsson, S. Flavobacteria blooms in four eutrophic lakes: linking population dynamics of freshwater bacterioplankton to resource availability. *Appl. Environ. Microbiol.* **73**, 3511–3518 (2007).

207. Tezuka, Y. Cation-dependent flocculation in a Flavobacterium species predominant in activated sludge. *Appl. Microbiol.* **17**, 222–226 (1969).

208. Świątczak, P. & Cydzik-Kwiatkowska, A. Performance and microbial characteristics of biomass in a full-scale aerobic granular sludge wastewater treatment plant. *Environ. Sci. Pollut. Res.* **25**, 1655–1669 (2018).

209. Gómez-Consarnau, L. *et al.* Light stimulates growth of proteorhodopsin-containing marine

Flavobacteria. *Nature* **445**, 210–213 (2007).

210. Yoshizawa, S. *et al.* Functional characterization of flavobacteria rhodopsins reveals a unique class of light-driven chloride pump in bacteria. *Proc. Natl. Acad. Sci.* **111**, 6732–6737 (2014).

APPENDICES

Appendix A: Supplemental works

FULL-LENGTH 16S rRNA GENE SEQUENCES FROM RAW SEWAGE SAMPLES SPANNING GEOGRAPHIC AND SEASONAL GRADIENTS IN CONVEYANCE SYSTEMS ACROSS THE UNITED STATES

Published in:

American Society for Microbiology: Resource Announcements

DOI: https://doi.org/10.1128/mra.00319-22

Emily Lou LaMartina[1], Angela L. Schmoldt[2], Ryan J. Newton[1]*

[1] School of Freshwater Sciences, University of Wisconsin-Milwaukee, 600 E Greenfield Ave., Milwaukee, WI 53204

[2] Great Lakes Genomics Center, Milwaukee, WI, USA

* Corresponding Author: newtonr@uwm.edu

(*Published 21 June 2022*)

Abstract

Wastewater microbiome research often relies on sequencing of hypervariable regions of 16S rRNA genes, which are difficult to classify at refined taxonomic levels. Here, we introduce a data set of near-full-length 16S rRNA genes from samples designed to capture known geographic and seasonal variations in municipal wastewater microbial communities.

Announcement

Wastewater-based monitoring for disease-causing entities is growing as a public health tool[195,196]. However, there remain significant gaps in understanding the inherent biology of sewage conveyance and its potential influence on monitoring efforts. To aid the characterization of wastewater microorganisms, 46 raw wastewater treatment plant (WWTP) influent samples underwent near full-length 16S rRNA gene sequencing. We selected samples that, according to previous work, encompass microbial community variability across geographic and seasonal gradients[4,159]. Raw influent (25-mL) was filtered onto 0.2-µm mixed cellulose ester filters (Sigma Millipore, #WHA10401770) from which DNA was extracted (FastDNA Spin Kit for Soil (MP Biomedicals, #116560200-CF) as described previously[4,159]. Genes were amplified using KAPA HiFi HotStart ReadyMix (Roche, #KK2602) under the following thermocycler conditions: 95°C for 5 min; 20 cycles of 98°C for 20 sec, 55°C for 45 sec, and 72°C for 3 min; 72°C for 5 min and with the primers 27F (5'- AGRGTTYGATYMTGGCTCAG-3') and 1492R (5'- RGYTACCTTGTTACGACTT). Each primer contained a pad sequence (GGTAG) followed by a unique 16-bp barcode (www.pacb.com) appended to the 5'-end. Prior to PCR, the barcoded

primers were phosphorylated with a T4 polynucleotide kinase (NEB, #M0201S) and ATP (NEB, #P0756S).

Following PCR, amplicons were equimolarly pooled and purified with AmPure PB beads (PacBio, #100-265-900). Libraries were created using Pacific Biosciences SMRTbell Express Template 2.0 (PacBio, #101-685-400) following the manufacturer's protocol. Amplicons were enzymatically repaired and ligated to a PacBio adapter to form the SMRTbell template. Templates were sequenced on a Sequel II System using Sequencing Primer v.4 (PacBio, #101-359-000) and the Sequel II 2.1 binding kit (PacBio, # 101-820-500). UW-Milwaukee Great Lakes Genomics Center (RRID:SCR_017838) provided PacBio sequencing services.

Default parameters were used for all software unless otherwise specified. BAM files from the PacBio Sequel II were converted to FASTQs with BEDtools v.2.30.0[197]. SeqKit v.2.2.0[163] was used to demultiplex FASTQs into individual files according to their unique barcodes. Primers were removed from demultiplexed files with Cutadapt[132]. Following a PacBio-specific protocol, DADA2 v.1.16 [133] on Galaxy v.22.01[198] was used to quality filter (max N = 0, max EE = 2), correct errors, and assign taxonomy with Silva v.138 [199] as a reference database. In most reads, the first 10 primer bases on the 3' end of the read were not trimmed by Cutadapt. These bases were removed in an exact-match approach (grep/cut). Resulting amplicon sequence variants (ASVs) were clustered to operational taxonomic units (OTUs) at 99.5% similarity with mothur v.1.43.0[200] and its protocol (https://mothur.org/wiki/cluster/).

Before demultiplexing, the raw FASTQ file had 7,750,870 reads, which condensed to 1,041 ASVs and 698 OTUs. A summary of raw, ASV, and OTU data is listed in Table 11. ASVs ranged from 1,383 to 1,553 base pairs, with a mean length of 1,455 base pairs. All ASVs were classified as bacteria and included 22 phyla, 35 classes, 71 orders, 116 families, 190 genera, and 158 species. The improved taxonomic resolution from full-length gene sequences resulted in 643 (61.8%) ASVs classified to species, compared to 3.48% in a V4-V5 hypervariable region study of a similar sample set[159].

Results summary

The most common OTUs are distinct between datasets. Communities expected to have a "warm" assemblage (according to a previous study[159]), such as those from the South US, are very different from "cold" communities (Fig. 16). Entire genera such as *Trichococcus* were completely absent from the most warm-like samples. In contrast, *Pseudomonas mendocina* were exclusively found in South US wastewater. Within- (richness) and between- (similarity) sample diversity tracks what has been shown in previous studies (Fig 17). Short-read V4-V5 analyses showed more within-sample diversity, however, long-read full 16S rRNA genes captured 96% of the short-read ASVs. Therefore, both short- and long-read analyses are sufficient for community analysis, but short-read (Illumina) data might better capture rare organisms, while long-read (PacBio) offers greater taxonomic resolution. As seen previously[159], wastewater temperature is a strong driver of microbial community structures (Fig. 18). Warm-like wastewater samples cluster apart from cold ones. Further, relative abundances of OTUs fluctuate according to those temperatures.

Data availability

Demultiplexed FASTQ files can be found on the NCBI Sequence Read Archive under BioProject PRJNA809416. Annotated files, additional analyses, and code are available on NewtonLabUWM GitHub. More information can be found at its website (https://loulanomics.github.io).

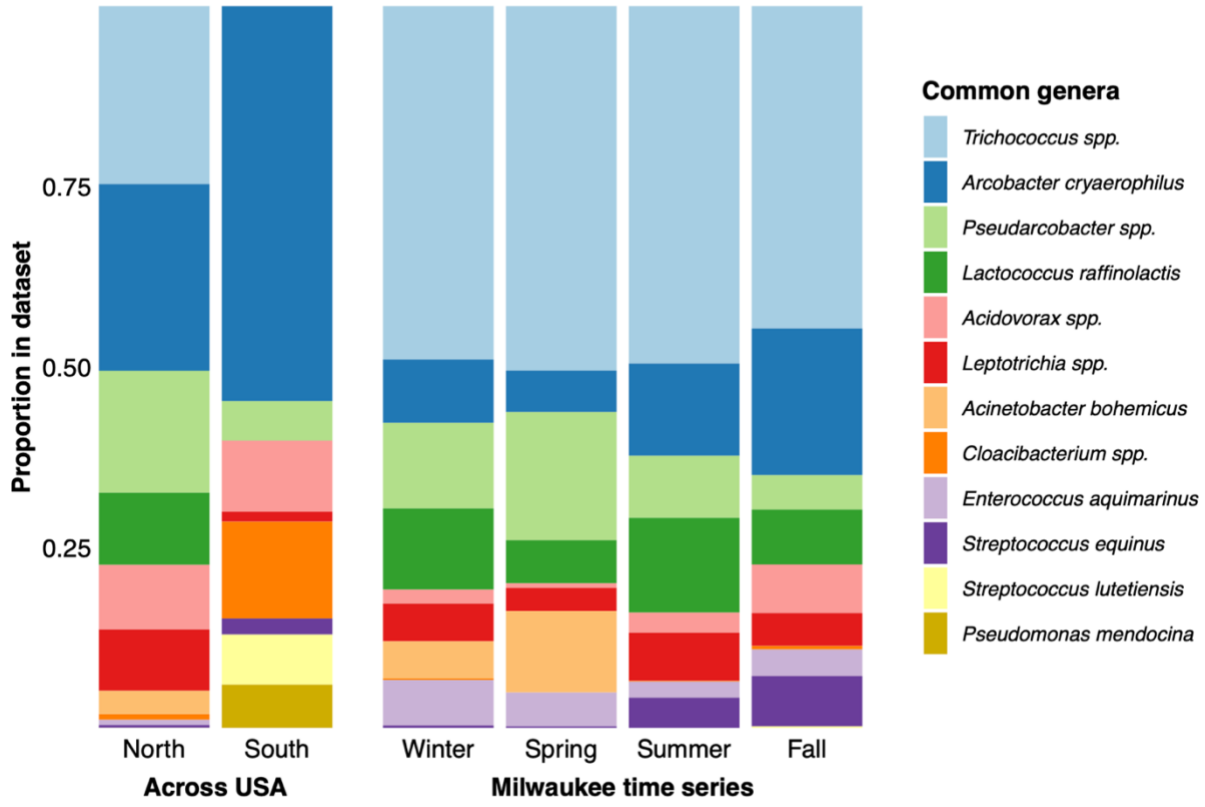

Figure 16. Wastewater genera using full-length 16S rRNA gene sequencing.

The top 5 most abundant OTUs in each wastewater sample set (north and south United States and winter, spring, summer, and fall in Milwaukee, Wisconsin) were identified, comprising 64.0% of all sequences. Bar height indicates the proportion of that OTU among the common OTUs analyzed. Bar colors denote genus and species assignments of OTUs.

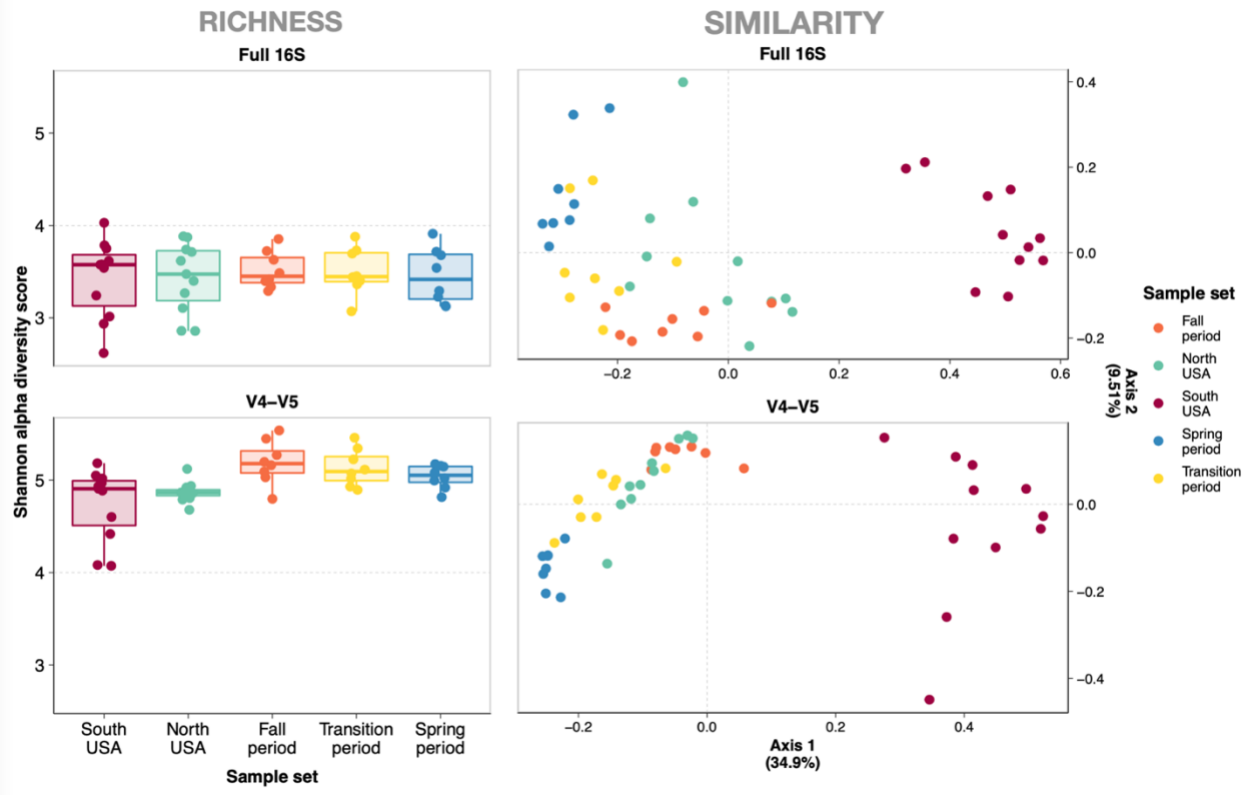

Figure 17. Community analysis using full-length 16S rRNA gene sequencing.

(left) Alpha diversity (richness) was calculated with the Shannon diversity index and (right) beta diversity (dis/similarity) was calculated with Bray-Curtis distance metrics. Diversity was measured in wastewater treatment plant influent from Milwaukee collected in Spring (blue point; Feb, Mar, Apr, May), Fall (orange points; Aug, Sep, Oct, Nov), and in between (transition period; yellow points; Dec, Jan, Jun, Jul) over two years (2016-2017; Table 1). Diversity of influent from northern US cities (green points) and southern US cities (purple points) were also measured (Table 2). Diversity was compared between datasets analyzing different amplicons, (bottom) the hypervariable V4-V5 region of 16S rRNA and (top) near full-length 16S rRNA genes. Boxes depict the median and first and third quartiles. Whisker lines extend to interquartile ranges x1.5 and points are outlier values.

Figure 18. Community clustering based on full-length 16S rRNA gene sequencing.

Dendrograms and heatmaps of top 5 most abundant OTUs per dataset (spring, fall, north, south) displayed at the phylum level in wastewater treatment plant influent from Milwaukee collected in over two years (2016-2017), northern US cities, and southern US cities.

Table 11. Summary of demultiplexed sequencing data (BioProject PRJNA809416).

BioSample [1]	Run [2]	State	Date	Raw reads [3]	ASVs [4]	OTUs [5]
SAMN26027580	SRR18111974	Montana	2013-01-16	8,124	29	25
SAMN26027581	SRR18111973	Oregon	2013-01-16	50,475	66	38
SAMN26027582	SRR18111962	Washington	2013-01-15	39,857	86	63
SAMN26027583	SRR18111951	Iowa	2013-01-15	48,102	73	57
SAMN26027584	SRR18111940	Nebraska	2013-01-23	49,510	63	39
SAMN26027585	SRR18111933	Wisconsin	2013-01-27	61,322	90	58
SAMN26027586	SRR18111932	Alaska	2013-01-23	44,505	89	65
SAMN26027587	SRR18111931	Wyoming	2013-01-24	40,151	37	21
SAMN26027588	SRR18111930	Colorado	2013-01-23	63,240	74	44
SAMN26027589	SRR18111929	Texas	2012-08-30	50,787	116	98
SAMN26027590	SRR18111972	Alabama	2012-08-14	38,960	70	55
SAMN26027591	SRR18111971	Georgia	2012-08-16	48,628	78	60
SAMN26027592	SRR18111970	California	2012-08-14	43,504	78	59
SAMN26027593	SRR18111969	Florida	2012-08-08	42,993	87	78
SAMN26027594	SRR18111968	Tennessee	2012-08-15	40,164	66	50
SAMN26027595	SRR18111967	Texas	2012-08-15	40,220	62	49
SAMN26027596	SRR18111966	Arizona	2012-08-15	39,140	62	51
SAMN26027597	SRR18111965	California	2012-08-21	48,423	59	36
SAMN26027598	SRR18111964	Florida	2012-08-21	43,786	151	109
SAMN26027599	SRR18111963	Hawaii	2012-09-07	43,509	57	42
SAMN26027600	SRR18111961	Minnesota	2013-01-16	52,709	58	34
SAMN26027601	SRR18111960	Ohio	2013-01-17	36,980	55	40
SAMN26027602	SRR18111959	Wisconsin	2016-04-07	37,955	83	61
SAMN26027603	SRR18111958	Wisconsin	2017-04-03	44,272	116	84
SAMN26027604	SRR18111957	Wisconsin	2016-08-03	37,141	52	41
SAMN26027605	SRR18111956	Wisconsin	2017-08-22	66,278	84	52
SAMN26027606	SRR18111955	Wisconsin	2016-12-07	53,471	66	43
SAMN26027607	SRR18111954	Wisconsin	2017-12-01	40,778	71	49
SAMN26027608	SRR18111953	Wisconsin	2016-02-08	99,767	95	61
SAMN26027609	SRR18111952	Wisconsin	2017-02-06	42,116	65	45
SAMN26027610	SRR18111950	Wisconsin	2016-01-06	72,965	110	74
SAMN26027611	SRR18111949	Wisconsin	2017-01-05	32,035	71	53
SAMN26027612	SRR18111948	Wisconsin	2016-07-18	46,548	55	40

BioSample [1]	Run [2]	State	Date	Raw reads [3]	ASVs [4]	OTUs [5]
SAMN26027613	SRR18111947	Wisconsin	2017-07-12	52,143	81	56
SAMN26027614	SRR18111946	Wisconsin	2016-06-08	48,861	55	41
SAMN26027615	SRR18111945	Wisconsin	2017-06-07	45,176	72	46
SAMN26027616	SRR18111944	Wisconsin	2016-03-02	29,799	47	28
SAMN26027617	SRR18111943	Wisconsin	2017-03-01	60,041	69	40
SAMN26027618	SRR18111942	Wisconsin	2016-05-02	40,253	83	64
SAMN26027619	SRR18111941	Wisconsin	2017-05-01	9,838	46	38
SAMN26027620	SRR18111939	Wisconsin	2016-11-03	38,702	58	39
SAMN26027621	SRR18111938	Wisconsin	2017-11-02	49,693	71	45
SAMN26027622	SRR18111937	Wisconsin	2016-10-05	50,908	71	55
SAMN26027623	SRR18111936	Wisconsin	2017-10-04	47,985	59	37
SAMN26027624	SRR18111935	Wisconsin	2016-09-21	36,116	59	49
SAMN26027625	SRR18111934	Wisconsin	2017-09-26	41,013	80	59

[1] BioSample ID on NCBI.
[2] Run ID on NCBI.
[3] Number of reads generated by PacBio Sequel II after demultiplexing.
[4] Number of amplicon sequence variants (ASVs) after processing with DADA2.
[5] Number of operational taxonomic units (OTUs) after clustering to 99.5% similarity with mothur.

Appendix B: Works in progress

PANGENOMIC EVALUATION OF SEWAGE FLAVOBACTERIA

Background

The genus *Flavobacterium* is extremely diverse and its members are adaptable to many different aquatic environments. *Flavobacterium* belong to the phylum *Bacteroidetes*, which have a unique protein secretion system, type IX, that is involved in virulence, polymer degradation, motility, and biofilm formation[201]. They are important fish pathogens due to their ability to secrete virulence factors and degrade fish tissues[202]. They are helpful for controlling algae blooms, as they flourish when algae die off, breakdown the cellulose, providing simpler sugars for the food web[203–206]. *Flavobacterium* use a gliding motility that allows them to crawl across surfaces, by secreting adhesins to the surface that rotate length-wise like a corkscrew, and can reverse direction, flip, pivot, and rotate[201]. Their biofilms aid WWTPs flocculation, during which suspended solids are clumped together and removed more easily[207]. Flocs also remove metals and phosphates from the water, protecting the ecosystems where effluent is discharged, and greatly improving wastewater treatment plant (WWTP) efficiency[208]. Incredibly, some supplement their chemotrophic lifestyle with phototrophy, using a distinct yellow-orange pigments that allow them to use light for energy[209,210]. Phototrophy gives *Flavobacterium* an incredible advantage by surviving periods when nutrients are limited.

Plan

Genetic differences of *Flavobacterium* across the urban water cycle will be investigated in a pangenome analysis. Full genomes of *Flavobacterium* were isolated from wastewater in September 2019 and February 2020. From a public genome database (https://gtdb.ecogenomic.org), genomes of *Flavobacterium* isolated from environments linked to the urban water cycle, such as fresh waterways, wild fish, pipes, and WWTPs, will be compiled. Interest is in *how* different they are in comparison, whether very genetically similar or distinct, as well as *what* makes them different, such as metabolism, virulence, pigments, and genome size. Because sewers are relatively new ecosystems, microbial populations within them must have come from natural systems, and a genetic analysis could improve understanding of microbial adaptability in urban landscapes.

Preliminary results

Out of 48 isolates from wastewater collected in September 2019 and February 2020 (Table 12), 19 were positively PCR-screened (Fig. 20) with *Flavobacterium* primers (Table 4).

Materials and methods

Creating a wastewater isolate catalogue

1. 1-L influent from Jones Island WWTP in Milwaukee, WI, was collected on 9/11/19 and 2/4/20.
2. 10-ml was reserved, the rest was autoclaved with 15-g agar, and poured into 60-mm petri plates. Once cooled, 100-µl of reserved wastewater was spread onto plates.

3. Plates were wrapped in parafilm and incubated at room temperature (22°C) in a BSL-2 lab for two weeks. During incubation, the room would experience brief periods of light, but for the most park was in complete darkness
4. Colony morphology was recorded for up to three isolates on each plate.
5. Colonies were then picked from plates with a sterile loop and inoculated into culture tubes with 2-ml 10x-diluted R2A (Table 13). Tubes were shaken in an incubator at 25°C for three days.
6. 100-µl of grown-up colonies were transferred to Cryovials with 1.4-ml freezing medium (Table 14) and stored at -80°C.

Reviving candidate Flavobacterium isolates

1. Colonies that were pigmented during initial isolation were selected as *Flavobacterium* candidates (Table 12 and Fig. 19).
2. Cryovials of candidates were pulled from -80°C storage in September 2022, after 36 and 31 months for the September 2019 and February 2020 isolates, respectively.
3. A flame-sterile loop was dipped into still-frozen cultures, spread onto 25-mm petri plates with 1x R2A, and streaked for CFUs.
4. Plates were stored in a 25°C incubator for two days. Photos of plates can be seen here.
5. Following, a single isolated colony was picked from each plate, inoculated into 2-ml 1x R2A, and shaken in an incubator at 25°C for three days.
6. 1-ml of grown-up cultures were boiled in 2-ml screw-cap tubes for 10 mins and centrifuged at top speed for 15 mins. 500-µl of supernatant with DNA was transferred to DNA elution tubes and stored indefinitely at -20°C. Cultures were stored at 4°C until ready for revival.

Pre-screening candidate isolates

1. DNA from boiled isolates were added to PCR assays with custom *Flavobacterium* primers[159] (Table 4).
2. Reactions were set up to 12-µl as follows: 6-µl GoTaq Green, 0.35-µl of 100-µM flavo11_sew_629F, 0.35-µl of 100-µM flavo11_sew_859R, 0.35-µl of 100-µM flavo42_sew_630F, 0.35-µl of 100-µM flavo42_sew_859R, 3.6-µl water, and 1-µl isolate DNA.
3. The following thermocycler conditions were used: 1 cycle at 95°C for n minutes; 40 cycles at 65°C for 1 minutes 1 cycle at 72°C for 2 minutes.
4. PCR products were added to a 3% agarose gel stained with 5% ethidium bromide and electrophoresed at 100V for 25 minutes alongside a 50-bp molecular ladder (Fig. 20).

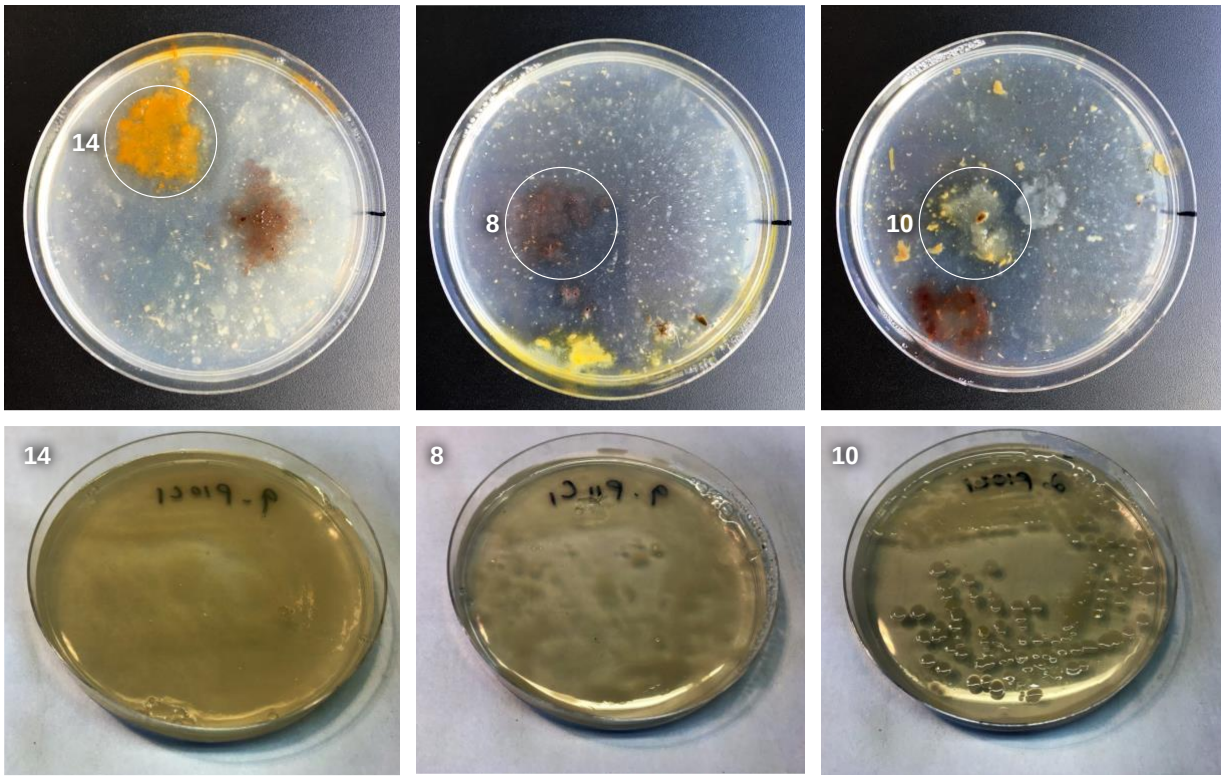

Figure 19. Petri plates growing *Flavobacterium*.

(top) Initial growth of wastewater on wastewater-derived medium. (bottom) Colonies restreaked on R2A.

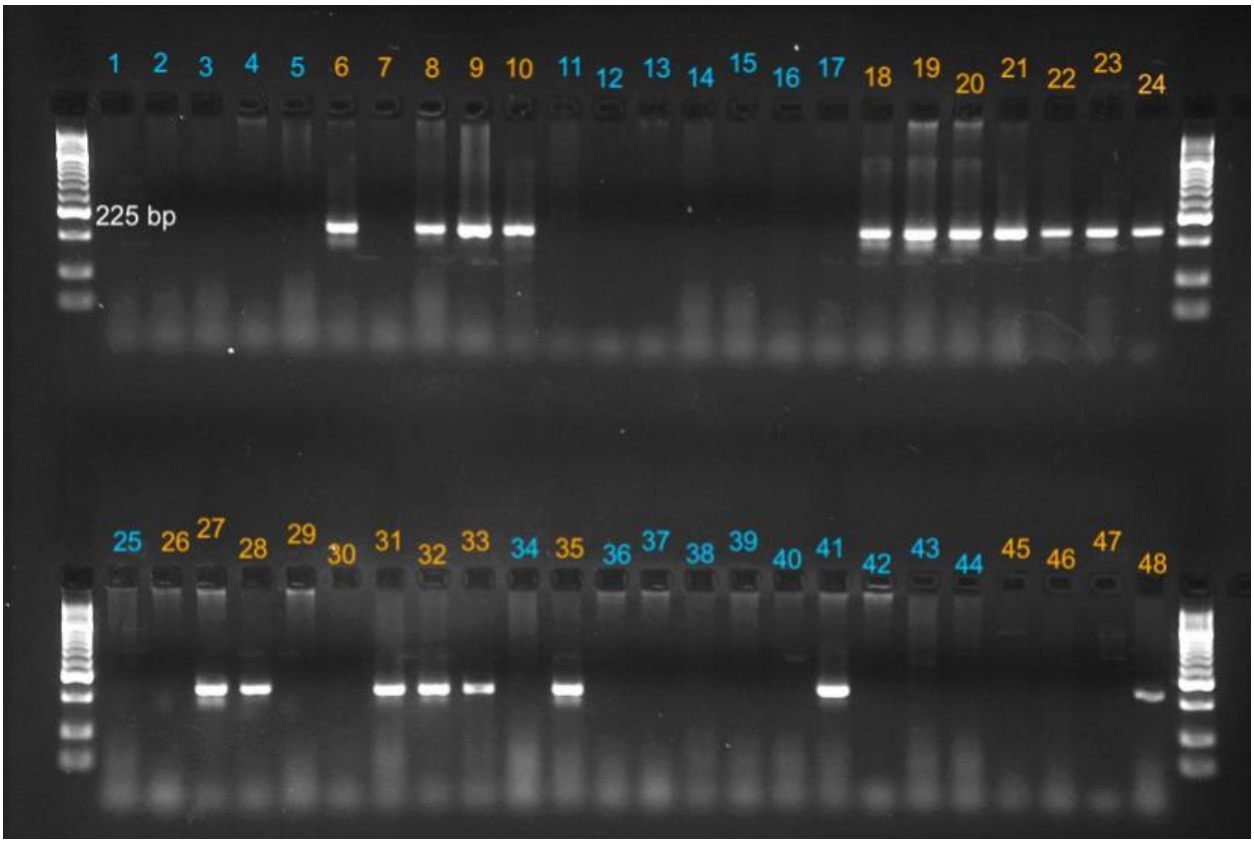

Figure 20. Gel results screening sewage isolates for *Flavobacterium*.

First and last wells on top and bottom rows were 50-bp ladder. Numbers indicate isolate number described in Table 12. Colors of number denote if culture was isolated September 2019 (orange) or February 2020 (blue).

Table 12. *Flavobacterium* isolate screening results.

Isolate information		Isolate morphology					Dates handled					PCR results [2]
No.	**ID [1]**	**Size**	**Shape**	**Color**	**Clarity**	**Growth**	**Collected**	**Isolated**	**Frozen**	**Revived**	**Cultured**	**Screen**
1	P03C1_021720	small	round	white	translucent	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
2	P26C1_021720	med	round	dark yellow	clear shiny	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
3	P04C1_021720	*N/A*	*N/A*	clear	clear	wet lawn	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
4	P16C2_021720	med	blob	dark yellow	clear	cluster	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
5	P16C1_021720	med	blob	dark yellow	clear	cluster	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
6	P12C2_091119	large	round	purple	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
7	P11C2_091119	large	round	orange	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	
8	P11C1_091119	large	round	purple	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
9	P10C2_091119	nickel	round	purple	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
10	P10C1_091119	quarter	round	orange	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
11	P27C2_021720	med	round	dark yellow	clear shiny	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
12	P09C2_021720	med	round	white	translucent	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
13	P09C1_021720	med	round	white	translucent	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
14	P10C1_021720	med	round	dark yellow	clear shiny	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
15	P13C1_021720	large	round	white	translucent	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
16	P12C1_021720	nickel	blob	dark yellow	clear	smear	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
17	P12C1_021720	nickel	blob	dark yellow	clear	smear	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
18	P18C1_091119	large	round	green-yellow	opaque	web	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
19	P18C1_091119	large	round	green-yellow	opaque	web	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
20	P22C1_091119	large	*N/A*	purple	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
21	P21C1_091119	large	*N/A*	purple	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
22	P21C1_091119	large	*N/A*	purple	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
23	P19C1_091119	large	round	purple	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
24	P18C2_091119	large	round	green-yellow	opaque	web	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
25	P28C1_021720	med	round	dark yellow	clear shiny	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	

Isolate information		Isolate morphology					Dates handled					PCR results [2]
No.	**ID [1]**	**Size**	**Shape**	**Color**	**Clarity**	**Growth**	**Collected**	**Isolated**	**Frozen**	**Revived**	**Cultured**	**Screen**
26	P17C3_091119	small	round	yellow	matte	single	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	
27	P27C1_091119	small	round	yellow	opaque	single	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
28	P25C2_091119	*N/A*	round	yellow	translucent shiny	single	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
29	P23C2_091119	small	round	yellow	matte	single	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	
30	P07C1_091119	med	round	yellowish	clear	single	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	
31	P24C2_091119	small	round	yellow	opaque	single	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
32	P06C2_091119	1 cm	long	yellowish	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
33	P16C1_091119	large	round	purple	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
34	P01C2_021720	med	blob	dark yellow	clear	cluster	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
35	P17C1_091119	large	*N/A*	purple	opaque	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+
36	P30C1_021720	med	round	dark yellow	clear shiny	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
37	P23C1_021720	med	round	dark yellow	clear shiny	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
38	P03C2_021720	small	round	white	translucent	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
39	P24C2_021720	med	round	dark yellow	clear shiny	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
40	P08C1_021720	pinprick	round	dark yellow	clear	cluster	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
41	P01C1_021720	med	blob	dark yellow	clear	cluster	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	+
42	P29C1_021720	large	round	white	translucent	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
43	P07C1_021720	med	round	dark yellow	clear shiny	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
44	P20C1_021720	small	round	white	translucent	single	2/17/20	3/2/20	3/5/20	9/8/22	9/10/22	
45	P05C1_091119	nickel	round	bright yellow	opaque shiny	cluster	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	
46	P01C2_091119	small	round	yellow	matte	single	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	
47	P12C3_091119	large	spreading	light yellow	opaque	spreading	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	
48	P01C1_091119	med	round	light yellow	shiny opaque	single	9/11/19	9/24/19	9/28/29	9/8/22	9/10/22	+

[1] P indicates plate number and C indicates colony number on that plate.

[2] Mixture of primers targeting *Flavobacteria* ASVs 11 and 42 in PCR and agarose gel electrophoresed. See Table 4 and Fig. 20.

Table 13. Growth medium recipe for *Flavobacterium* cultures.

Ingredient	Amount
Yeast Extract	0.5 g
Peptone	0.5 g
Casamino Acids	0.5 g
Glucose (Dextrose)	0.5 g
Soluble starch	0.5 g
Sodium pyruvate	0.3 g
K_2HPO_4	0.3 g
$MgSO_4$ x $7H_2O$	0.05 g

Instructions

1. Add all ingredients, minus agar, to a large autoclave-safe container. Bring volume to 1-L with deionized water and mix thoroughly.
2. Adjust to pH 7.2 ± 0.2 (HCl if too high, NaOH if too low) at room temperature.
3. If making petri plates, add 15-g agar or gellan gum (Gelzan).
4. Autoclave for >15 min at 15 psi and 121°C.
5. If making petri plates: pour as soon as possible while still liquid.
6. If dispensing to culture tubes: let cool.

Table 14. Recipe for long-term isolate storage.

Ingredient	Volume (ml)
Liquid growth medium	127
100% glycerol	11
Dimethyl sulfoxide (DMSO)	11

Instructions

1. Autoclave growth medium and glycerol in separate containers.
2. Filter-sterilize DMSO with syringe and 0.22-µm filter.
3. To make freezing medium, mix growth medium, DMSO, and glycerol.
4. Aliquot 1.4-ml freezing medium to Cryovials.
5. Add 100-µl microbial culture and store at -80°C.

Appendix C: Collaborations

Publications

Status	Citation
Published	Burch, T. R., Newton, R. J., Kimbell, L. K., **LaMartina, E. L.**, O'Malley, K., Thomson, S. M., ... & McNamara, P. J. (2022). Targeting current and future threats: recent methodological trends in environmental antimicrobial resistance research and their relationships to risk assessment. *Environmental Science: Water Research & Technology (IF: 4.3*).
Published	Kimbell, L. K., **LaMartina, E. L.**, Kappell, A. D., Huo, J., Wang, Y., Newton, R. J., & McNamara, P. J. (2021). Cast iron drinking water pipe biofilms support diverse microbial communities containing antibiotic resistance genes, metal resistance genes, and class 1 integrons. *Environmental Science: Water Research & Technology (IF 4.3)*, 7(3), 584-598.
In review	Scarim, G., **LaMartina, E. L.**, Venkiteshwaran, K., Zitomer, D., Newton, R. J., & McNamara, P. J. A novel method to quantify activated sludge foaming and its use to elucidate the role of microbial community structure.
In review	MacLellan-Hurd, R.A., **LaMartina, E.L.**, Liao, Q., Troy, C., & Bootsma, H. Quagga mussel (*Dreissena rostriformis bugensis*) biodeposit effects on benthic nutrient cycling in Lake Michigan. Target: *Journal of Great Lakes Research (IF: 2.48)*.
In review	Jones, D. C., Lewis J., **LaMartina, E. L.**, Dahl, A. J., Holavanahalli, N. N., Newton, R. J., & Skwor, T.A. One Health and Global Health View of Antimicrobial Susceptibility through the "Eye" of *Aeromonas*: Systematic Review and Meta-Analysis. *International Journal of Antimicrobial Agents* (IF 15.4).
In review	Kimbell, L. K., Kohls, S., Thomson, S. M., **LaMartina, E. L.**, Wang, Y., Newton, R. J., & McNamara, P. J. Corrosion Inhibitors Influence Antibiotic Resistance, Metal Resistance, and Microbial Communities in a Source for Drinking Water.
Dissertation	Brennaman, A., **LaMartina, E. L.**, Newton, R. J., & Richards, P. Tales from the Tooth Worm: Health and the Historic Oral Microbiome at the Milwaukee County Poor Farm Cemetery.
In prep.	Kennedy, E., **LaMartina, E. L.**, Kappell, A. D., Harrison, K., Mayer, B., Zitomer, D., Newton, R. J., & McNamara, P. J. On microbial community and dewaterability in anaerobically digested primary and Bio-P biosolids.

Research projects

Group	Description
Ryan Newton UW-Milwaukee	DNA and cDNA sequencing of full-length 16S rRNA genes from groundwater
Kazuaki Matsui Kindai University	Microbial community analysis in a weekly, one-year time series of raw wastewater from two treatment plants using quantitative standards
Dong-Fang Deng UW-Milwaukee	Microbial community analyses of yellow perch and sturgeon guts undergoing diet augmentations
Todd Miller UW-Milwaukee	Quantifying *Cyanobacteria* and associated toxin-producing genes in an impacted lagoon using droplet digital PCR (ddPCR)
Jhonatan Sepulveda UW-Milwaukee	Microbial community analysis in rhizoplane and rhizosphere of aquaponic plants

Mentorships

Student	Description and awards
Erica Kallas Nicolet High School	Effect of added concentrations of *Melaleuca alternifolia* oil on the growth of *Pseudomonas fluorescens* in Lake Michigan Water Stockholm Junior Water Prize National Oceanic and Atmospheric Administration Award
Garrett Scapellato UW-Milwaukee	*Cloacibacterium*: developing a model system to study wastewater conveyance. Summer Undergraduate Research Award
Lauren Bonofiglio UW-Milwaukee	Who are Milwaukee Microbes? Summer Undergraduate Research Award