Environmental Surveillance of Enteric Pathogens in Zimbabwe Urban Wastewater Using Metagenomics

Samantha Yan Cheung

A thesis

submitted in partial fulfillment of the

requirements for the degree of:

Master of Science

University of Washington

2024

Committee:

John Scott Meschke

Karen Levy

Nicolette Zhou

Program Authorized to Offer Degree:

Environmental and Occupational Health Sciences

University of Washington

Abstract

Environmental Surveillance of Enteric Pathogens in Zimbabwe Urban Wastewater Using Metagenomics

Samantha Yan Cheung

Chair of the Supervisory Committee:

John Scott Meschke

Department of Environmental and Occupational Health Sciences

High diarrheal disease burden remains an urgent concern in low to middle income countries, greatly affecting children under the age of 5 years old and those living with HIV and AIDS. The capital city Harare, Zimbabwe, has been at the center of recurrent typhoid and cholera outbreaks in the last few decades. Treatment of these diarrheal illnesses and other infectious diseases has also become increasingly difficult with the rapid rise of antimicrobial resistance in the country. Despite these major concerns, there is little understanding about the enteric pathogens and antimicrobial resistance that contribute to disease burden in Harare. Clinical surveillance provides advantages in characterizing select pathogens and their characteristics from patients, but this method is expensive, resource-intensive, and highly dependent on health-seeking behaviors and accessibility, providing limited understanding about pathogen distribution and community impacts. Environmental surveillance of wastewater can supplement these gaps because all residents on a sewage system contributes to the wastewater, providing simple, composite

samples that can give an improved understanding about both pathogens and antimicrobial resistance in the community. This study evaluated the effectiveness of environmental surveillance with shotgun metagenomics as a tool to characterize a wide range of enteric pathogens, antibiotic resistance genes (ARGs), and-virulence factor genes (VFGs) in wastewater from Harare, Zimbabwe. Between April 19 and May 9, 2019, wastewater samples from three high density, low-income suburbs and three low density, high-income suburbs were collected and processed for next-generation sequencing (n=18). Diversity analyses of the cumulative metagenomes revealed a significant difference in alpha diversity (Shannon-Wiener diversity index) and no significant difference in beta diversity (Bray-Curtis dissimilarity distance) between samples from high-income suburbs and low-income suburbs. The top enteric pathogens detected in all wastewater samples include *Arcobacter spp.* and *Aeromonas spp.*, common bacteria found in environmental water and associated with mild to moderate gastroenteritis. Pathogens of high antimicrobial resistance (AMR) and clinical concern, including *Staphylococcus aureus, Pseudomonas aeruginosa,* and *Salmonella enterica*, were detected in all wastewater samples. From 510 predicted ARGs, we identified 139 unique ARGs across all samples. Of the total number of predicted ARGs, 66 were chromosomal and 212 were plasmid, suggesting the majority of ARGs are on mobile elements. Out of 3580 curated genes within 1381 VFs in the Virulence Factor Database, we detected 412 genes within 50 VFs in the samples. The top three virulence factor function classes were delivery, adherence, and motility, which play a major role in toxin secretion, host cell colonization, immune modulation, and cell survival. The findings provide a foundation for future studies to explore the potential of environmental surveillance and shotgun metagenomics as a public health monitoring tool for enteric disease.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Diarrheal Disease

Diarrheal diseases account for a large percentage of the overall disease burden worldwide, especially in children under five years of age in low-middle income countries (LMICs). They lead to approximately 525,000 child deaths worldwide annually (*World Health Organization, 2017*), while antimicrobial resistance (AMR) related illnesses directly cause approximately 1.25 million deaths (children and adult) annually *(WHO, 2019*). Diarrheal disease is also compounded by the high HIV infection rate in LMICs, leading to more vulnerable individuals in the population due to compromised immune systems (*UNICEF, 2023*).

With the exception of Poliovirus*,* pathogen and antimicrobial resistance distribution in the environment are not well understood in areas with recurring outbreaks of diarrheal disease. Hygiene studies and evaluation of water, sanitation, and hygiene (WASH) intervention impacts are heavily concentrated in Bangladesh, India, and Kenya with over 50 studies each, compared to many other LMICs with fewer than 10 studies or no data (*Chirgwin et.al, 2021*). Throughout Africa, there is an uneven distribution of studies on AMR in wastewater, and even within countries—research on AMR from wastewater treatment plants was concentrated in the KwaZulu-Natal and Eastern Cape provinces of South Africa *(Abia et al, 2023*). Insufficient data of WASH services in countries with high burden of diarrheal disease impedes decision-making for improved policies, recommendations, and interventions.

1.2 Harare, Zimbabwe

1.2.1 Wastewater System and Water Use

Harare is the capital city of Zimbabwe with a population of 1.8 million people (*Chirisa et al., 2017*). Residential areas are categorized based on density—high, medium, and low (*Nhapi, 2004*). High-density suburbs are primarily composed of low-income residents living in small housing stands around 300 square meters; medium-density suburbs contain stands around 1000 square meters, and low-density suburbs contain stands around 2000 square meters for high-income residents (*Nhapi, 2004*).

Two major wastewater treatment plants serve around 80% of the city population: Firle and Crowborough (*Chirisa et al., 2017*). The Firle plant is located in the outskirts of Glenview, and the Crowborough plant is located between Kuwadzana and Mufakose (Supplementary Figure 1). In addition to these plants, two waste stabilization ponds in Marlborough and Donnybrook, along with an aeration plant in Hatcliffe, treat wastewater influent (*Nhapi, 2004*). About 20% of the population are not accommodated by these systems, which are informal settlements with widespread pit latrines and open defecation practices (*Chirisia et al., 2017*).

This wastewater system was originally built to only serve a population of 250,000 people, about one-sixth of the current population. All facilities suffer from overloading and poor maintenance, leading to the poor treatment and discharge of effluent into pastures and the Mukusivi and Marimba Rivers, which flow into Lakes Chivero and Manyame, the city's main sources of water (*Nhapi, 2004*). Activated sludge treatment at Firle and Crowborough attempt to remove high

phosphorus and nitrogen levels, but both plants do not adequately remove pathogens from the effluent, further polluting the water sources (*Nhapi, 2004*). In addition, the lack of maintenance has led to low-income, high-density suburbs continuously experiencing burst sewage pipes and leaking sewage, but the Harare City Council have continued to fail in addressing these issues due to financial constraints (*Tsiko & Togarepi, 2012*). For the same financial reasons, the Harare City Council has also not been able to repair damaged boreholes. As a result, drinking water from these boreholes is easily contaminated, increasing the risk of diarrheal infection for low-income residents dependent on these sources of poor-quality water.

1.2.2 Diarrheal Disease Burden

Zimbabwe is an LMIC affected by a large diarrheal disease burden. In 2014, there were an estimated 763,136 outpatient diarrheal cases in Zimbabwe (*Zimbabwe MOHCC 2014*), with an annual incidence of 25.1 per 1,000 persons in Harare province. In 2017-2018, a typhoid outbreak began in a Harare suburb (*N'cho 2019*), resulting in 3,187 suspected cases. Following the typhoid outbreak, a large cholera outbreak with 8,535 cases occurred in 2018, (*WHO | Cholera – Zimbabwe, 2018*) leading to an oral cholera vaccine campaign (*WHO | Zimbabwe to vaccinate 1.4 million people against cholera in Harare, 2018*). Prior to the 2018 outbreak, there was a cholera outbreak in 2008 to 2009 that led to more than 4000 deaths. The most recent cholera outbreak began in February 2023 outbreak and has persisted for over a year, spreading from Chegutu town to all 10 provinces with 1649 suspected cases, according to an International Federation of Red Cross report on June 1st (*IFRC, 2023*).

Despite recurring outbreaks, pathogen and antimicrobial resistance distribution in the environment are not well understood in Zimbabwe. Hygiene studies and evaluation of water, sanitation, and hygiene (WASH) intervention impacts are heavily concentrated in Bangladesh, India, and Kenya with over 50 studies each, compared to many other LMICs with fewer than 10 studies or no data (*Chirgwin et.al, 2021*). Throughout Africa, there is an uneven distribution of studies on antimicrobial resistance (AMR) in wastewater, even within countries—AMR research is concentrated in the KwaZulu-Natal and Eastern Cape provinces of South Africa *(Abia et al, 2023*). Insufficient data on the impact of WASH services on environmental distribution in countries with high burden of diarrheal disease, such as Zimbabwe, impedes decision-making for improved policies, recommendations, and interventions.

The lack of extensive WASH research in Zimbabwe may partially explain the short-term success of implemented interventions to waterborne illnesses During the October 2016 typhoid outbreak in Harare, interventions included borehole and sewage line repairment, installation of inline chlorinators, and distribution of basic sanitation and water purification items. Despite thorough risk factor identification and response efforts, resurgence occurred the next year in October 2017 (*CDC, 2018*). Unsuccessful long-term resolution of diarrheal pathogens could be partially attributed to gaps in current knowledge regarding the characteristics and distribution of diarrheal pathogens in Harare. New information through a comprehensive approach with environmental surveillance may initiate further studies to understand disease burden and population impacts with environmental surveillance.

1.3 Antimicrobial Resistance

Antimicrobial resistance is a rapidly growing global public health concern, contributing to an estimated 5 million deaths worldwide in 2019 *(CDC, 2022*). Driven by misuse in agriculture,

medicine, and animal husbandry, AMR is widespread in both non-pathogenic and pathogenic bacteria, as horizontal gene transfer of ARGs can occur between an extensive diversity of species (*Rout et.al, 2023*). AMR reduces treatment efficacy against a growing number of infectious diseases (*WHO, 2023*). Reducing AMR requires overcoming gaps in understanding about the movement of pathogens and ARGs through the community with more high-quality research in LMICs.

Zimbabwe has observed a rise in antimicrobial resistance. The main driver of increasing AMR burden is antimicrobial overuse, which has only grown during the COVID-19 pandemic. As part of the country's COVID-19 management guidelines, prescription of ceftriaxone and azithromycin for antibiotic therapy for symptomatic cases was recommended (*Adebisi et.al, 2021; Chitungo et al., 2022*). Inappropriate approaches to treatment most likely stem from limited access to laboratory diagnostics, lack of antibiotic prescription training for health professionals, and dependence on a syndromic approach for treatment (*Olaru et.al, 2022*). In addition, unregulated counterfeit antibiotics in street markets are easily accessible to the local residents, and they are inexpensive compared to medications provided in pharmacies.

Analyzing clinical isolates can provide a glimpse of AMR trends in a subset of the population, but sampling can be invasive for participants. In 2017, Zimbabwe began to move towards addressing AMR with a One Health AMR surveillance system. As part of these efforts, the World Health Organization (WHO) Extended Spectrum Beta-Lactamase Escherichia coli (ESBL Ec) Tricycle protocol was implemented into the surveillance system between March 2021 to 2022. The protocol was evaluated in a study that compared *E. coli* isolates collected from pregnant women and poultry. Preliminary results revealed high levels of ESBL Ec were found in two-thirds of processed poultry products and in over a quarter of rectal swabs from pregnant women in labor, suggesting highly drug-resistant bacteria are circulating in Zimbabwe (*Chimbwanda, 2022*). Although the study was informative about antibiotic resistance spread, sampling was highly invasive and completed on women in labor, a highly stressful and vulnerable experience. Collection of wastewater for environmental surveillance does not require invasive human sampling or deidentification, while providing an understanding about AMR trends in the sampled community.

1.4 Virulence Factors

Virulence factors are another set of characteristics significant to understanding pathogen proliferation and disease severity. Virulence factors are defined as molecules that assist the bacterium with colonization of the host at the cellular level and enhance the potential to cause disease (*Sharma A. et.al, 2017; Liu et al., 2022*). Virulent strains within a species are observed to carry virulence factor genes, while non-pathogenic strains within the same species typically do not carry virulence factor genes or express them.

Environmental factors can influence virulence factor expression, decreasing or increasing expression that aid in survival among harsh and rapidly changing conditions, while affecting the pathogen's potential to colonize the host and cause disease. One example is *Vibrio cholerae*: in the aquatic environment, the bacterium forms biofilms, stores nutrients, and slows metabolism to survive numerous environmental stressors; in the host, *V. cholerae* transitions to expression of virulence factors that overcome the host gastrointestinal environment and protection (*Conner, 2016*). Detection of virulence factor genes can provide insight on the adaptive mechanisms of

bacterial pathogens in environmental survival and virulence, while understanding their functions can reveal clues about the environment of the community (*Rout et.al, 2023*).

1.5 Environmental Surveillance

Environmental surveillance is a form of active surveillance that can reveal pathogen circulation in communities, validating findings from clinical surveillance. Although clinical surveillance is the gold standard for disease monitoring, it is a form of passive surveillance that is dependent on medium- to high-complexity laboratory infrastructure and intensive resources unavailable in many areas of LMICs, while providing an incomplete representation of disease burden in the communities. In addition, clinical surveillance solely relies on health-seeking behaviors or access from patients, leading to detection bias towards pathogens that cause severe symptoms or additional complications that require clinical care. This dependence on clinical cases results in low case reporting, missed sub-clinical circulation, and limited sampling of the population. Wastewater surveillance can supplement clinical surveillance to provide a more accurate estimate of disease burden, detect outbreaks earlier, monitor antimicrobial resistance, and assess the effectiveness of control programs and infrastructure.

Environmental surveillance has proven to be effective in the Global Polio Eradication Initiative efforts by monitoring transmission of polio within communities through detection of poliovirus circulation in sewage water, complementing clinical surveillance findings (*Asghar et al., 2014*). Development of early warning systems for poliovirus, norovirus, and hepatitis also arose from the successful use of wastewater tracking for these pathogens (*Street et.al, 2020*). During the COVID-19 pandemic, wastewater surveillance garnered immense attention for widespread usage to track SARS-CoV-2 in communities and international travel. Following the peak of the pandemic, there have been rapidly growing discussions to incorporate wastewater surveillance for additional notifiable diseases, such as gastrointestinal illnesses caused by *Salmonella* and *Campylobacter* species, into new or existing national infectious disease surveillance systems.

1.6 Shotgun Metagenomics

Previous studies utilized culture, standard PCR, and qPCR as common detection methods for pathogens. These cost-effective techniques are best for the detection of a few targets of interest, but they are inadequate for understanding the environmental microbiome. With inexpensive, new advancements in sequencing technology, next-generation sequencing can provide a wide range of information about numerous microorganisms present in a single wastewater sample, including taxonomic classification, antimicrobial resistance, virulence factors, and additional characteristics that aid in its survival and proliferation *(Rout et.al, 2023*). Shotgun metagenomic sequencing comprehensively samples most DNA present compared to 16S rRNA sequencing, which targets a single gene for bacterial identification alone. This allows for the identification of VFGs and ARGs, adding significant value to the identification of microorganisms present in an environmental sample. With a dataset of 757 shotgun metagenomics sewage samples from a longitudinal global study, studies were able to successfully identify over 55,000 ARGs and characterize geographical and temporal patterns of AMR, as well as examine origin and patterns of sewage bacterial organisms (*Munk et al., 2022, Jespersen et al., 2023*). Analysis linking potential pathogens, ARGs, and VFGs together can build a comprehensive picture about enteric diseases in communities.

1.7 Purpose

This pilot study was conducted in Harare, Zimbabwe, to characterize enteric pathogens, antimicrobial resistance, and virulence factors present in wastewater of Harare, Zimbabwe. Environmental surveillance paired with shotgun metagenomic sequencing will provide extensive information about the characteristics and disease burden of a wide range of diarrheal pathogens in the community environment. The findings will establish new baseline data on the pathogens present in urban wastewater of Harare. This provides greater insight as to what pathogens may play a role in overall disease burden. Applying shotgun metagenomics to wastewater surveillance can broaden the detectable range of pathogens while simultaneously tracking AMR and identifying significant virulence factors, allowing us to examine the utility of these paired techniques as a population health monitoring tool.

We predicted that we would identify common enteric bacterial pathogens containing virulence factors genes associated with gastrointestinal disease and relevant antimicrobial resistance genes. This study also explored the association between neighborhood socioeconomic status and the diversity of enteric pathogens, antimicrobial resistance genes, and virulence factor genes in wastewater. Demonstrating the potential of environmental surveillance paired with shotgun metagenomic sequencing to provide a wide range of relevant information about enteric pathogens can guide tentative implementation of routine wastewater surveillance of target diarrheal diseases, which is crucial to reduce diarrheal disease burden.

CHAPTER 2: THE STUDY

2.1 Abstract

High diarrheal disease burden remains an urgent concern in low to middle income countries, greatly affecting children under the age of 5 years old and those living with HIV and AIDS. The capital city Harare, Zimbabwe, has been at the center of recurrent typhoid and cholera outbreaks in the last few decades. Treatment of these diarrheal illnesses and other infectious diseases has also become increasingly difficult with the rapid rise of antimicrobial resistance in the country. Despite these major concerns, there is little understanding about the enteric pathogens and antimicrobial resistance that contribute to disease burden in Harare. Clinical surveillance provides advantages in characterizing select pathogens and their characteristics from patients, but this method is expensive, resource-intensive, and highly dependent on health-seeking behaviors and accessibility, providing limited understanding about pathogen distribution and community impacts. Environmental surveillance of wastewater can supplement these gaps because all residents on a sewage system contributes to the wastewater, providing simple, composite samples that can give an improved understanding about both pathogens and antimicrobial resistance in the community. This study evaluated the effectiveness of environmental surveillance with shotgun metagenomics as a tool to characterize a wide range of enteric pathogens, antibiotic resistance genes (ARGs), and-virulence factor genes (VFGs) in wastewater from Harare, Zimbabwe. Between April 19 and May 9, 2019, wastewater samples from three high density, low-income suburbs and three low density, high-income suburbs were collected and processed for next-generation sequencing (n=18). Diversity analyses of the cumulative metagenomes revealed a significant difference in alpha diversity (Shannon-Wiener diversity index) and no significant difference in beta diversity (Bray-Curtis dissimilarity distance) between samples from high-income suburbs and low-income suburbs. The top enteric pathogens detected in all wastewater samples include *Arcobacter spp.* and *Aeromonas spp.*, common bacteria found in environmental water and associated with mild to moderate gastroenteritis. Pathogens of high antimicrobial resistance (AMR) and clinical concern, including *Staphylococcus aureus, Pseudomonas aeruginosa,* and *Salmonella enterica*, were detected in all wastewater samples. From 510 predicted ARGs, we identified 139 unique ARGs across all samples. Of the total number of predicted ARGs, 66 were chromosomal and 212 were plasmid, suggesting the majority of ARGs are on mobile elements. Out of 3580 curated genes within 1381 VFs in the Virulence Factor Database, we detected 412 genes within 50 VFs in the samples. The top three virulence factor function classes were delivery, adherence, and motility, which play a major role in toxin secretion, host cell colonization, immune modulation, and cell survival. The findings provide a foundation for future studies to explore the potential of environmental surveillance and shotgun metagenomics as a public health monitoring tool for enteric disease.

Keywords: Antimicrobial resistance, enteric pathogens, environmental surveillance, shotgun metagenomics, wastewater, Zimbabwe

2.2 Introduction

Diarrheal diseases account for a large percentage of the overall disease burden worldwide, especially in children under five years of age in low-middle income countries (LMICs). They lead to approximately 525,000 child deaths worldwide annually (*World Health Organization, 2017*), while antimicrobial resistance (AMR) related illnesses directly cause approximately 1.25 million deaths (children and adult) annually *(WHO, 2019*). Diarrheal disease is also compounded by the high HIV infection rate in LMICs, leading to more vulnerable individuals in the population due to compromised immune systems (*UNICEF, 2023*).

The distribution of pathogens and antimicrobial resistance in the environment is not well understood in areas with recurring outbreaks of diarrheal disease. Hygiene studies and evaluation of water, sanitation, and hygiene (WASH) intervention impacts are heavily concentrated in Bangladesh, India, and Kenya with over 50 studies each, compared to many other LMICs with fewer than 10 studies or no data (*Chirgwin et.al, 2021*). Throughout Africa, there is an uneven distribution of studies on AMR in wastewater, and even within countries—research on AMR from wastewater treatment plants was concentrated in the KwaZulu-Natal and Eastern Cape provinces of South Africa *(Abia et al, 2023*). Insufficient data of WASH services in countries with high burden of diarrheal disease impedes decision-making for improved policies, recommendations, and interventions.

Zimbabwe is an LMIC affected by a large diarrheal disease burden. In 2014, there were an estimated 763,136 outpatient diarrheal cases in Zimbabwe (*Zimbabwe MOHCC 2014*), with an annual incidence of 25.1 per 1,000 persons in Harare province. In 2017-2018, a typhoid outbreak began in a Harare suburb (*N'cho 2019*), resulting in 3,187 suspected cases. Following the typhoid outbreak, a large cholera outbreak with 8,340 cases occurred in 2018, (*WHO | Cholera – Zimbabwe, 2018*) leading to an oral cholera vaccine campaign (*WHO | Zimbabwe to vaccinate 1.4 million people against cholera in Harare, 2018*). Prior to the 2018 outbreak, there was a cholera outbreak in 2008 to 2009 that led to more than 4000 deaths. The most recent cholera outbreak began in February 2023 outbreak and has persisted for over a year, spreading from Chegutu town to all 10 provinces with 1649 suspected cases, according to an International Federation of Red Cross report on June 1st (*IFRC, 2023*).

Despite recurring outbreaks, pathogen and antimicrobial resistance distribution in the environment are not well understood in Zimbabwe. Hygiene studies and evaluation of water, sanitation, and hygiene (WASH) intervention impacts are heavily concentrated in Bangladesh, India, and Kenya with over 50 studies each, compared to many other LMICs with fewer than 10 studies or no data (*Chirgwin et.al, 2021*). Throughout Africa, there is an uneven distribution of studies on antimicrobial resistance (AMR) in wastewater, even within countries—AMR research is concentrated in the KwaZulu-Natal and Eastern Cape provinces of South Africa *(Abia et al, 2023*). Insufficient data on the impact of WASH services on environmental distribution in countries with high burden of diarrheal disease, such as Zimbabwe, impedes decision-making for improved policies, recommendations, and interventions.

Harare is the capital city of Zimbabwe with a population of 1.8 million people (*Chirisa et al., 2017*). Residential areas are categorized based on density—high, medium, and low (*Nhapi, 2004*). High-density suburbs are primarily composed of low-income residents living in small

housing stands around 300 square meters; medium-density suburbs contain stands around 1000 square meters, and low-density suburbs contain stands around 2000 square meters for high-income residents (*Nhapi, 2004*).

Two major wastewater treatment plants serve around 80% of the city population: Firle and Crowborough (*Chirisa et al., 2017*). The Firle plant is located in the outskirts of Glenview, and the Crowborough plant is located between Kuwadzana and Mufakos (Supplementary Figure 1). In addition to these plants, two waste stabilization ponds in Marlborough and Donnybrook, along with an aeration plant in Hatcliffe, treat wastewater influent (*Nhapi, 2004*). About 20% of the population are not accommodated by these systems, which are informal settlements with widespread pit latrines and open defecation practices (*Chirisia et al., 2017*).

The wastewater system was originally built to only serve a population of 250,000 people, about one-sixth of the current population. All facilities suffer from overloading and poor maintenance, leading to the poor treatment and discharge of effluent into pastures and the Mukusivi and Marimba Rivers, which flow into Lakes Chivero and Manyame, the city's main sources of water (*Nhapi, 2004*). Activated sludge treatment at Firle and Crowborough attempt to remove high phosphorus and nitrogen levels, but both plants do not adequately remove pathogens from the effluent, further polluting the water sources (*Nhapi, 2004*). In addition, the lack of maintenance has led to high-density suburbs continuously experiencing burst sewage pipes and leaking sewage, but the Harare City Council have continued to fail in addressing these issues due to financial constraints (*Tsiko and Togarepi, 2012*). As a result, the poor access to safe drinking water may increase the distribution of pathogens in the environment and community.

Antimicrobial resistance is another global public health concern that has rapidly grown in urgency, contributing to an estimated 5 million deaths worldwide in 2019 *(CDC, 2022*). Driven by misuse in agriculture, medicine, and animal husbandry, AMR is widespread in both non-pathogenic and pathogenic bacteria, as horizontal gene transfer of ARGs can occur between an extensive diversity of species (*Rout et.al, 2023*). AMR reduces treatment efficacy against a growing number of infectious diseases (*WHO, 2023*). Reducing AMR requires overcoming gaps in understanding about the movement of pathogens and ARGs through the community with more high-quality research in LMICs.

Despite efforts in AMR monitoring and regulation, Zimbabwe continues to see a rise in antimicrobial resistance. The main driver of increasing AMR burden is antimicrobial overuse, which has only grown during the COVID-19 pandemic. As part of the country's COVID-19 management guidelines, prescription of ceftriaxone and azithromycin for antibiotic therapy for symptomatic cases was recommended (*Adebisi et.al, 2021; Chitungo et al., 2022*). Inappropriate approaches to treatment most likely stem from limited access to laboratory diagnostics, lack of antibiotic prescription training for health professionals, and dependence on a syndromic approach for treatment (*Olaru et.al, 2022*). In addition, unregulated counterfeit antibiotics in street markets are easily accessible to the local residents, and private pharmacies have a history of issuing medications without a prescription (*Gwatidzo et al., 2017*).

Virulence factors are not typically used to inform public health action, but they are a set of characteristics significant to understanding pathogen proliferation and disease severity. Virulence factors are defined as molecules that assist the bacterium with colonization of the host at the

cellular level and enhance the potential to cause disease (*Sharma A. et.al, 2017; Liu et al., 2022*). Virulent strains within a species are observed to carry virulence factor genes, while non-pathogenic strains within the same species typically do not carry virulence factor genes or express them.

Environmental factors can influence virulence factor expression, decreasing or increasing expression that aid in survival among harsh and rapidly changing conditions, while affecting the pathogen's potential to colonize the host and cause disease. One example is *Vibrio cholerae*: in the aquatic environment, the bacterium forms biofilms, stores nutrients, and slows metabolism to survive numerous environmental stressors; in the host, *V. cholerae* transitions to expression of virulence factors that overcome the host gastrointestinal environment and protection (*Conner, 2016*). Detection of virulence factor genes can provide insight on the adaptive mechanisms of bacterial pathogens in environmental survival and virulence, while understanding their functions can reveal clues about the environment of the community (*Rout et.al, 2023*).

Environmental surveillance is a form of active surveillance that can reveal pathogen circulation in communities, validating findings from clinical surveillance. Although clinical surveillance is the gold standard for disease monitoring, it is a form of passive surveillance that is dependent on medium- to high-complexity laboratory infrastructure and intensive resources unavailable in many areas of LMICs, while providing an incomplete representation of disease burden in the communities. In addition, clinical surveillance solely relies on health-seeking behaviors or access from patients, leading to detection bias towards pathogens that cause severe symptoms or additional complications that require clinical care. This dependence on clinical cases results in low case reporting, missed sub-clinical circulation, and limited sampling of the population. Wastewater surveillance can supplement clinical surveillance to provide a more accurate estimate of disease burden, detect outbreaks earlier, monitor antimicrobial resistance, and assess the effectiveness of control programs and infrastructure.

Environmental surveillance has proven to be effective in the Global Polio Eradication Initiative efforts by monitoring transmission of polio within communities through detection of poliovirus circulation in sewage water, complementing clinical surveillance findings (*Asghar et al., 2014*). Development of early warning systems for poliovirus, norovirus, and hepatitis also arose from the successful use of wastewater tracking for these pathogens (*Street et.al, 2020*). During the COVID-19 pandemic, wastewater surveillance garnered immense attention for widespread usage to track SARS-CoV-2 in communities and international travel. Following the peak of the pandemic, there have been rapidly growing discussions to incorporate wastewater surveillance for additional notifiable diseases, such as gastrointestinal illnesses caused by *Salmonella* and *Campylobacter* species, into new or existing national infectious disease surveillance systems.

Previous studies utilized culture, standard PCR, and qPCR as common detection methods for pathogens. These cost-effective techniques are best for the detection of a few targets of interest, but they are inadequate for understanding the environmental microbiome. With inexpensive, new advancements in sequencing technology, next-generation sequencing can provide a wide range of information about numerous microorganisms present in a single wastewater sample, including taxonomic classification, antimicrobial resistance, virulence factors, and additional characteristics that aid in its survival and proliferation *(Rout et.al, 2023*). Shotgun metagenomic sequencing comprehensively samples most DNA present compared to 16S rRNA sequencing,

which targets a single gene for bacterial identification alone. This allows for the identification of VFGs and ARGs, adding significant value to the identification of microorganisms present in an environmental sample. With a dataset of 757 shotgun metagenomics sewage samples from a longitudinal global study, studies were able to successfully identify over 55,000 ARGs and characterize geographical and temporal patterns of AMR, as well as examine origin and patterns of sewage bacterial organisms (*Munk et al., 2022, Jespersen et al., 2023*). Analysis linking potential pathogens, ARGs, and VFGs together can build a comprehensive picture about enteric diseases in communities.

This pilot study was conducted in Harare, Zimbabwe, to characterize enteric pathogens, antimicrobial resistance, and virulence factors present in wastewater of Harare, Zimbabwe. Environmental surveillance paired with shotgun metagenomic sequencing could provide extensive information about the characteristics and disease burden of a wide range of diarrheal pathogens in the community environment. The primary aim was to identify common enteric bacterial pathogens containing virulence factor genes associated with gastrointestinal disease and relevant antimicrobial resistance genes. This study also explored the association between neighborhood socioeconomic status and the diversity of enteric pathogens, antimicrobial resistance genes, and virulence factor genes in wastewater. Demonstrating the potential of environmental surveillance paired with shotgun metagenomic sequencing to provide a wide range of relevant information about enteric pathogens can guide tentative implementation of routine wastewater surveillance of target diarrheal diseases, which is crucial to reduce diarrheal disease burden.

2.3 Methods

The author did not participate in sample collection, processing, extraction, and sequencing.

2.3.1 Sample Collection and Filtration

The bag-mediated filtration system (BMFS v.2) was utilized to collect and process up to 6 litres of water from six suburbs in Harare, Zimbabwe (*Fagnant et al., 2017, 2018; Zhou et al., 2018; Ruhanya et al., 2022).* Suburbs were categorized as high-income, low-density (Marlborough, Northeastern, and Avonlea) or low-income, high-density (Budiriro, Budiriro West, and Shelter Zimbabwe) based on housing size. High-income, low-density suburbs contained housing stands around 2000 m^2. Low-income, high-density suburbs contained housing stands less than 300 m^2. Between April 19, 2019, and May 9, 2019, wastewater samples were collected at a manhole that feeds into a wastewater pump station in each suburb on three separate dates (n = 18). Sampling locations and wastewater pipe networks can be observed in Supplementary Figure 2. Volume, weather, and temperature of day were recorded at time of collection (Supplementary Table 1). The samples were filtered in the field, then transferred to the University of Zimbabwe Virology Laboratory. Filters were stored in a cooler on ice until reaching the University of Zimbabwe Virology Laboratory on the day of collection.

2.3.2 Sample Processing and Extraction

Sample processing and extraction was performed on all 18 wastewater samples at the University of Zimbabwe Virology Laboratory (Figure 1). From collection to shipment, chain of custody was properly conducted for all samples. At the University of Zimbabwe Virology Lab, elution, concentration, and extraction was conducted, then samples were shipped to the U.S. to complete sequencing and data analysis.

The BMFS Elution Device, v.2, was used to perform a double elution on each of the filter housing samples. Sample eluates underwent secondary concentration using a skim milk flocculation method. Elution and secondary concentration methods are detailed further in Supplementary Information.

Secondary concentrated samples were extracted with the Qiagen All Prep Power Fecal DNA/RNA Kit, with an input volume of 200 uL of the resuspended secondary concentration pellet and an elution volume of 60 uL. Extracts were frozen immediately at -20C in 0.5ml plastic tubes, prepped for shipment, then transported to the United States.

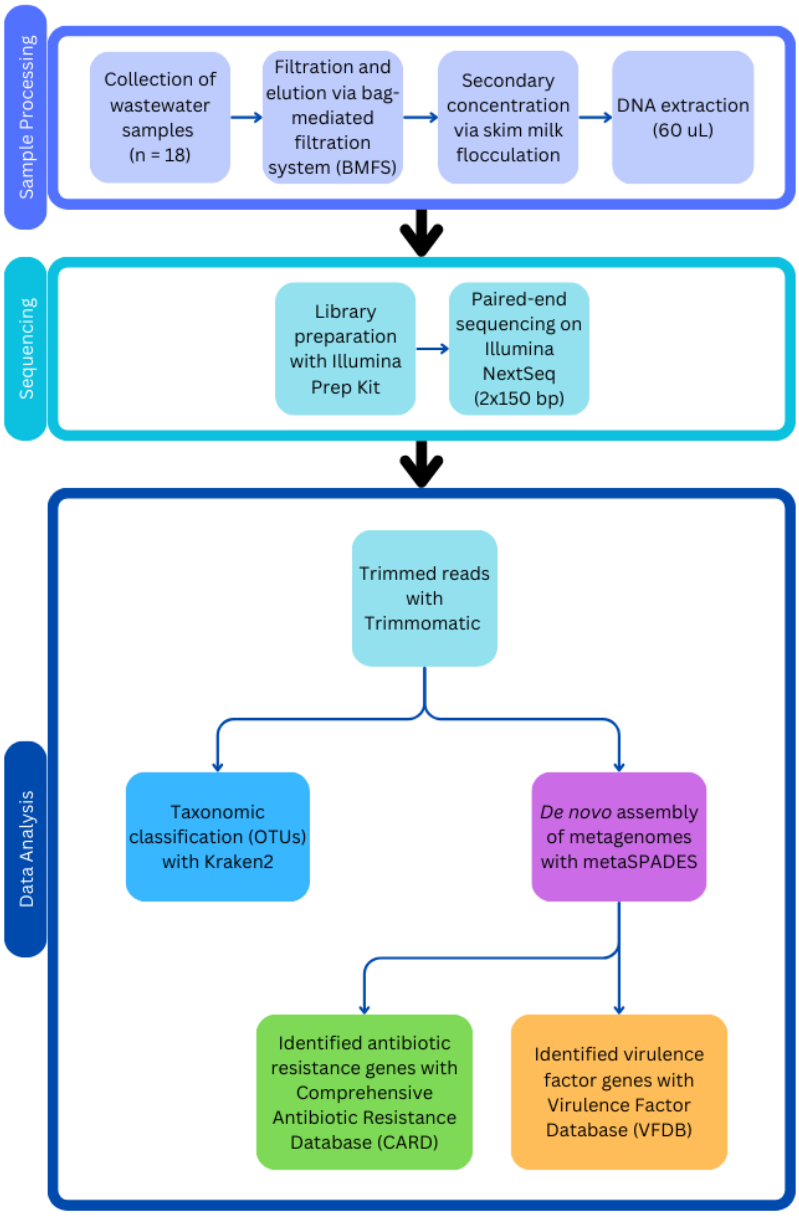


Figure 1. Flowchart of Study Design and Methods for Wastewater Samples: Simplified graphical flowchart of the methods

2.3.3 Shotgun Metagenomic Sequencing

DNA concentration of each sample was quantified with the Qubit Fluorometer and Qubit dsDNA HS Assay Kit (Supplementary Table 2). Extracts were prepared for sequencing using the Illumina DNA Prep kit, including tagmentation, a PCR protocol amplification, and wash steps. Paired end sequencing was conducted on an Illumina NextSeq System with an output of 2x150 bp paired-end reads per sample.

2.3.4 Metagenome Analysis

2.3.4.1 Quality Check and Pre-processing

Sequence quality was assessed with FastQC. The FASTA files of the samples were merged based on their respective suburb, then sequences were trimmed with Trimmomatic ver.0.40 (*Bolger, Lohse,* and *Usadel, 2014*) to remove adapter sequences.

Shotgun metagenomic sequencing generated a total of 143 million reads from all wastewater samples. Raw read count from each individual sample ranged from 2 million reads to 13.7 million reads. The range of alpha diversity scores of individual samples grouped by income status was much larger compared to the range of alpha diversity scores of combined sequences by suburb, grouped by income status. Median Shannon diversity scores of individual samples and merged samples were similar for both income statuses. However, the significant difference in Shannon diversity scores observed between high-income and low-income suburbs for merged reads was not present in the same comparison with individual sample reads. Beta diversity of individual samples and merged samples were also similar for both income statuses. Individual samples cluster tightly within neighborhoods of respective income status, with low-income and high-income clusters overlapping each other.

Although we acknowledge the difference in alpha diversity comparisons, the study lacked the relevant metadata to provide background context to explain this observation, as well as the widely variable number of reads in each sample. As the wastewater samples were collected in the same three-week period and suburb locations, we chose to merge the sequences of individual samples by respective suburb for downstream analyses, as observed in previous studies (*Rajeev et al., 2023; Ma et al., 2023*).

2.3.4.2 Databases

Trimmed paired sequences were *de novo* assembled using metaSPAdes ver.3.15.2 (*Nurk et al., 2017*), and assigned operational taxonomic units (OTUs) using Kraken2 ver.2.0.7 (*Wood, Lu,* and *Langmead, 2019*). Following OTU assignment by Kraken2, Bracken was utilized to determine relative abundance at the phylum and species level, as well as to identify potential bacterial pathogens (*Lu et al., 2017*). A comprehensive repository of bacterial pathogens of human clinical relevance globally was referenced for identification of bacterial pathogens (*Bartlett et al., 2022*). An additional list concerning antimicrobial resistant clinical isolates that cause AMR-related illnesses and deaths in Zimbabwe was referenced to identify major pathogens relevant in Zimbabwe (*Sartorius & Gray et al., 2023*).

The assembled metagenomes were submitted to the Comprehensive Antibiotic Resistance Database (CARD, ver.3.2.1) to identify ARGs with open reading frames (ORFs) predicted by Prodigal and analyzed by the Resistance Gene Identifier (*Hyatt et al., 2010; Alcock et al., 2023*). Output contained ARGs with perfect and strict hits. Perfect hits are defined as perfect matches to the reference sequences, while strict hits are unknown variants of the reference sequences that

meet the curated similarity threshold established by the RGI. The relative abundances of ARGs were visualized in a heatmap generated in R with the Complex Heatmap package (reference), displaying the top 50 most abundant genes. Most drug classes were based on the default CARD classification system, while "Multidrug" was defined by multiple drug classes to one ARG in the output. Plasflow was employed with default parameters to identify chromosome-associated and plasmid-associated ARGs for all ARG-like ORF contigs (*Krawzyck, Lipinski, Dziembowski, 2018*).

Assembled metagenomes were also submitted to the Virulence Factor Database (VFDB, curated VF-related genes, set A) to identify VFGs (Liu et al., 2022). Output was filtered by 90% identity and 70% query length coverage of the predicted protein. Best match was identified by selecting the identity with the smallest e-value (*Rodriguez & Konstantinidis, 2016*). Estimated average sequencing coverage was determined by assuming a Zero-Inflated Poisson (ZIP) distribution, and filtering out contigs with a ZIP greater than or equal to 0.3 (*Rodriguez & Knostantinidis, 2016*).

2.3.4.3 Diversity Tests

Alpha and beta diversity was determined for the assembled metagenomes, ARGs, and VFGs in the wastewater samples using the vegan package v.2.6-4 in R Studio. Detected counts of each genus, ARG, or VFG were inputted into the diversity analyses. Alpha diversity was calculated by the Shannon-Wiener diversity index. Scores were compared between high-income and low-income suburbs. Beta diversity was calculated by the Bray-Curtis dissimilarity distance to evaluate compositional differences between wastewater samples by socioeconomic status. Distances were visualized in non-metric multidimensional scaling (nMDS) plots.

2.3.4.4 Statistical Analysis

Statistical analyses were conducted with the stats and vegan packages in R. Significant differences in Shannon diversity scores between wastewater samples by socioeconomic status were assessed with the Wilcoxson signed-rank test. PERMANOVA was applied to the Bray-Curtis dissimilarity distances to assess for differences between wastewater samples from high-income suburbs and low-income suburbs. Statistical significance was determined by applying an alpha of 0.1 on p-values as the threshold.

2.4 Results

2.4.1 Relative abundance and diversity of microbial communities and bacterial pathogens in wastewater

Shotgun metagenomic sequencing generated a total of 143 million reads from all wastewater samples. Read count from each suburb ranged from 13.9 million reads to 29.2 million reads (Supplementary Table 3). A total of 11,972,209 contigs were classified into 60 phyla, 1978 genera, and 7183 species with Kraken2.

Alpha diversity was higher within samples from low-income neighborhoods compared to samples from high-income neighborhoods (Supplementary Figure 4A). We detected a significant difference in Shannon diversity scores between high-income and low-income suburbs for microbial composition ($p = 0.1$). In contrast to the significant difference in microbial diversity between high-income and low-income suburbs, Bray-Curtis dissimilarity between these

socioeconomic groups was not significant (Supplementary Figure 5A, p = 0.7). The microbial diversity is similar between high-income and low-income wastewater samples.

Microbial composition by phyla in wastewater samples was similar across suburbs (Supplementary Figure 3). The most dominant phylum in wastewater from each suburb was *Proteobacteria* (68.6% to 84.0%), followed by *Firmicutes* (4.9 % to 13%), *Bacteroidetes* (1.8% to 5.1%), and *Uroviricota* (0.9% to 6.4%). Increased abundance of *Firmicutes* and decreased abundance of *Proteobacteria* was observed in Northeastern and Budiriro neighborhoods compared to the other neighborhoods.

Microbial composition by genera in wastewater samples was also similar between suburbs (Figure 2). The top genera were *Arcobacter, Aeromonas,* and *Pseudomonas*. *Pseudomonas* is markedly increased and *Salmonella* is decreased in Avonlea compared to the relative abundance of these two genera in other suburbs.

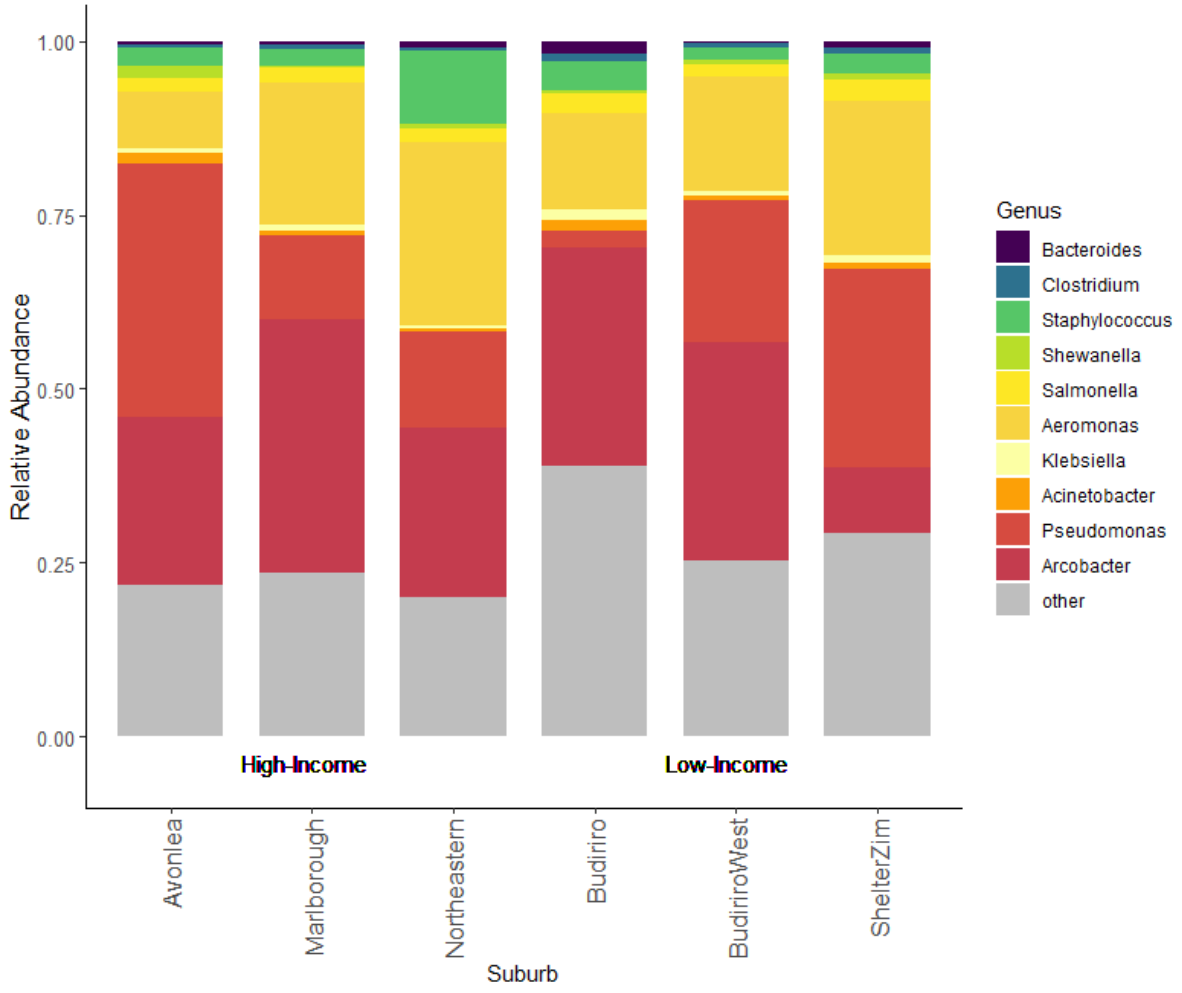


Figure 2. Relative Abundance of the Top 10 Genera in Wastewater Samples by Suburb: Genera not included in the top 10 were grouped into "Other".

The most abundant bacterial pathogens identified from all wastewater samples were *Arcobacter cryaerophilus*, *Aeromonas veronii*, *Aeromonas hydrophila*, and *Staphylococcus aureus* (Figure 3). This is consistent with the microbial composition observed at the genera level. *Salmonella*

enterica was also detected as part of the top 10 bacterial species, but the serotype was not identifiable. Bacterial composition was similar across suburbs and socioeconomic status, with variations in relative abundance of each species. An increased presence of *Aeromonas hydrophila* and a marked decrease of *A. cryaerophilus* was observed in Shelter Zimbabwe wastewater samples compared to other suburbs.

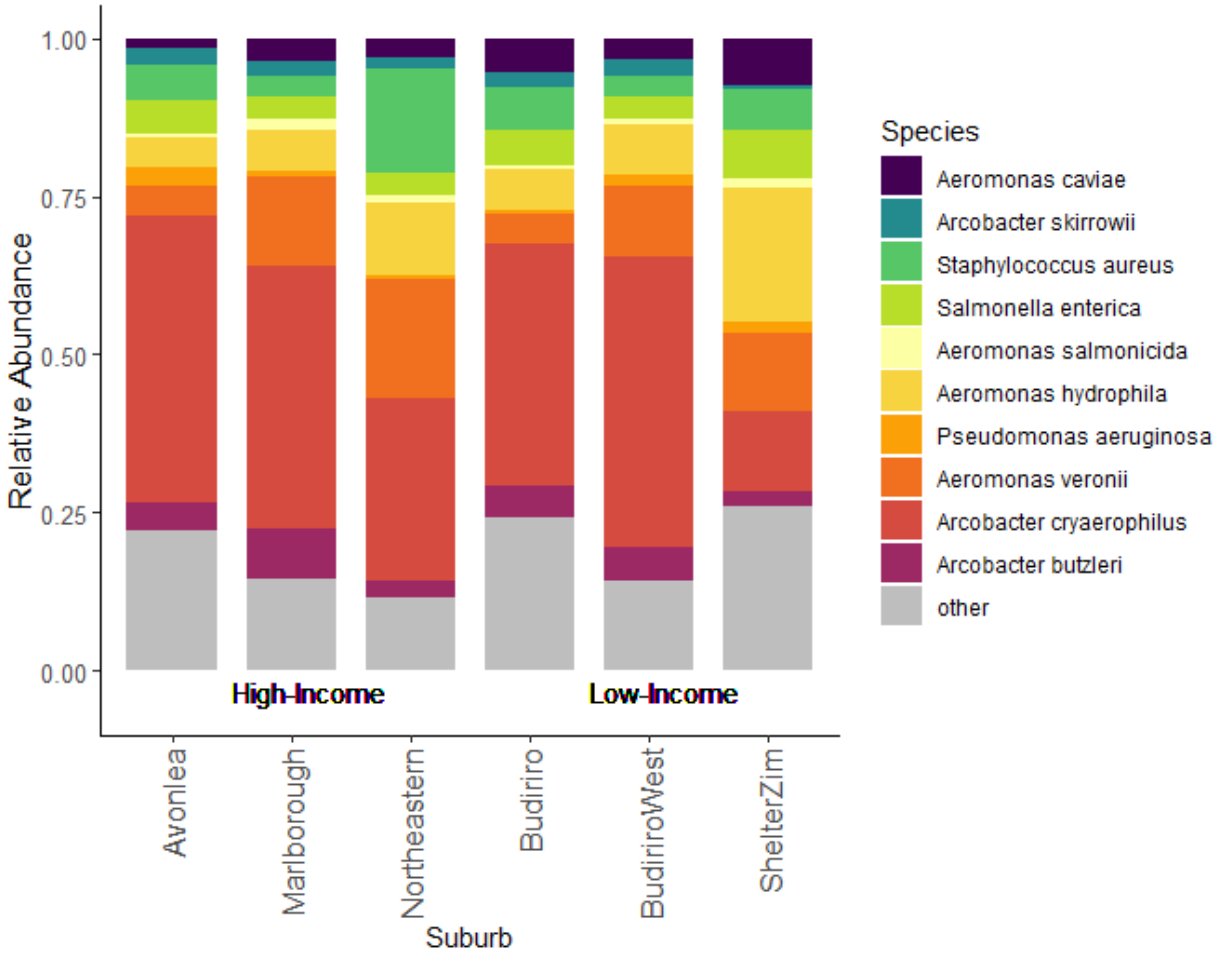


Figure 3. Relative Abundance of the Top 10 Bacterial Pathogens in Wastewater Samples by Suburb: Top 10 bacterial pathogenic species, defined by the *Bartlett et al., 2022* paper, are displayed. Species not included in the top 10 were grouped into "Other".

Pathogens of high AMR concern and commonly isolated from clinical cases in Zimbabwe were detected from wastewater (*Sartorius & Gray et al.,* 2023). *S. aureus*, *Pseudomonas aeruginosa*, and *Escherichia coli* were the top three most abundant pathogens across all suburbs (Figure 4). Pathogenic *E. coli* strains were not identifiable. This may be resolved with the detection of key virulence factors of pathogenic *E. coli* strains in the wastewater. Many *Aeromonas* and *Arcobacter* species observed in Figure 2 did not appear in Figure 3, as these two genera were not considered in the list of pathogens of AMR concern in Zimbabwe. Bacterial composition remained similar between neighborhoods and socioeconomic status, with slight variations in relative abundance of each species. A marked increase in *S. aureus* and a decrease in *E. coli* was observed in samples from the Northeastern neighborhood compared to other neighborhoods.

Under “Other”, an additional pathogen to note is *Shigella sp*., which was present in all neighborhoods at low relative abundance.

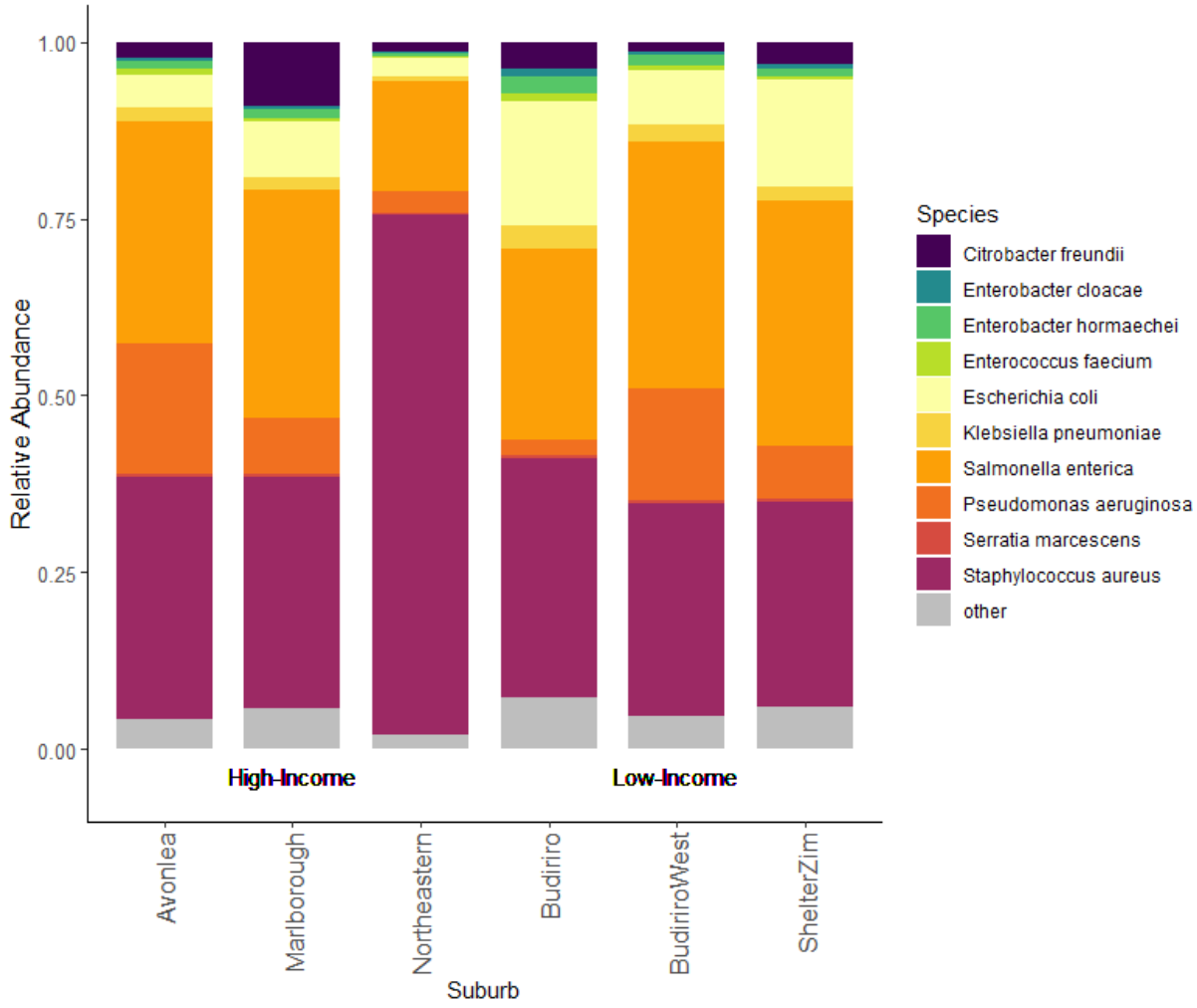


Figure 4. Relative Abundance of the Top 10 Bacterial Pathogens of High AMR and Clinical Concern: Stacked bar chart displaying top 10 bacterial pathogenic species of high AMR and clinical concern in Harare, defined by the *Sartorius & Gray et al., 2023* paper for Zimbabwe. Species not included in the top 10 were grouped into “Other”.

2.4.2 Antibacterial resistance genes

With the CARD RGI, we detected 139 unique ARGs from wastewater samples. Shannon diversity scores were similar within and between high-income and low-income neighborhoods (Supplementary Figure 4B, p = 1). The nMDS plot revealed high-income neighborhoods clustered together near the low-income neighborhoods Budiriro and Shelter Zimbabwe, while Budiriro West was most different from the other suburbs (Supplementary Figure 5B). Bray-Curtis dissimilarity distances were not significantly different between high-income and low-income neighborhoods (p = 0.5).

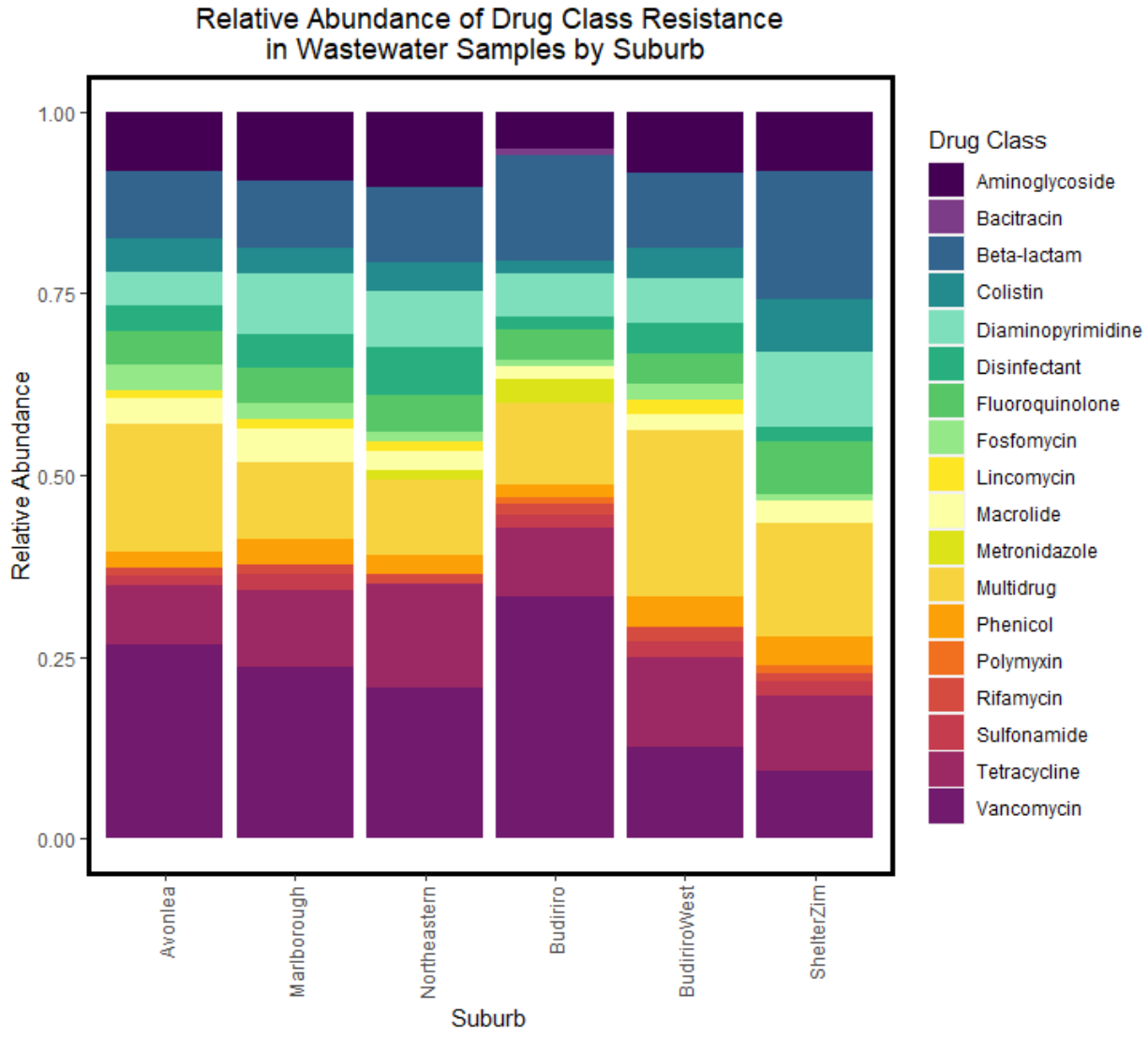


Figure 5. Relative Abundance of Drug Class Resistance by Suburb: Stacked bar chart displaying the relative abundance of ARGs categorized by drug class resistance.

We detected 510 ARG-like ORFs and 139 unique genes. Of the 10 resistance mechanisms included in CARD, we detected 5 mechanisms. Antibiotic inactivation was the most commonly detected mechanism in the ARG-like ORFs (26.9%), followed by target alteration (26.3%), and efflux (24.1%). Of the 61 drug classes in CARD, we identified 19 drug classes, with resistance to vancomycin (22.2%), beta-lactam (12.4%), and tetracycline antibiotics (10.6%) as the top three drug classes in the ARG-like ORFs (Figure 5). Multiple drug classes detected in a single contig

were classified as “multidrug”, constituting 14% of the ARG-like ORFs. There was low resistance to metronidazole, lincomycin, polymyxin B, and bacitracin (<1%).

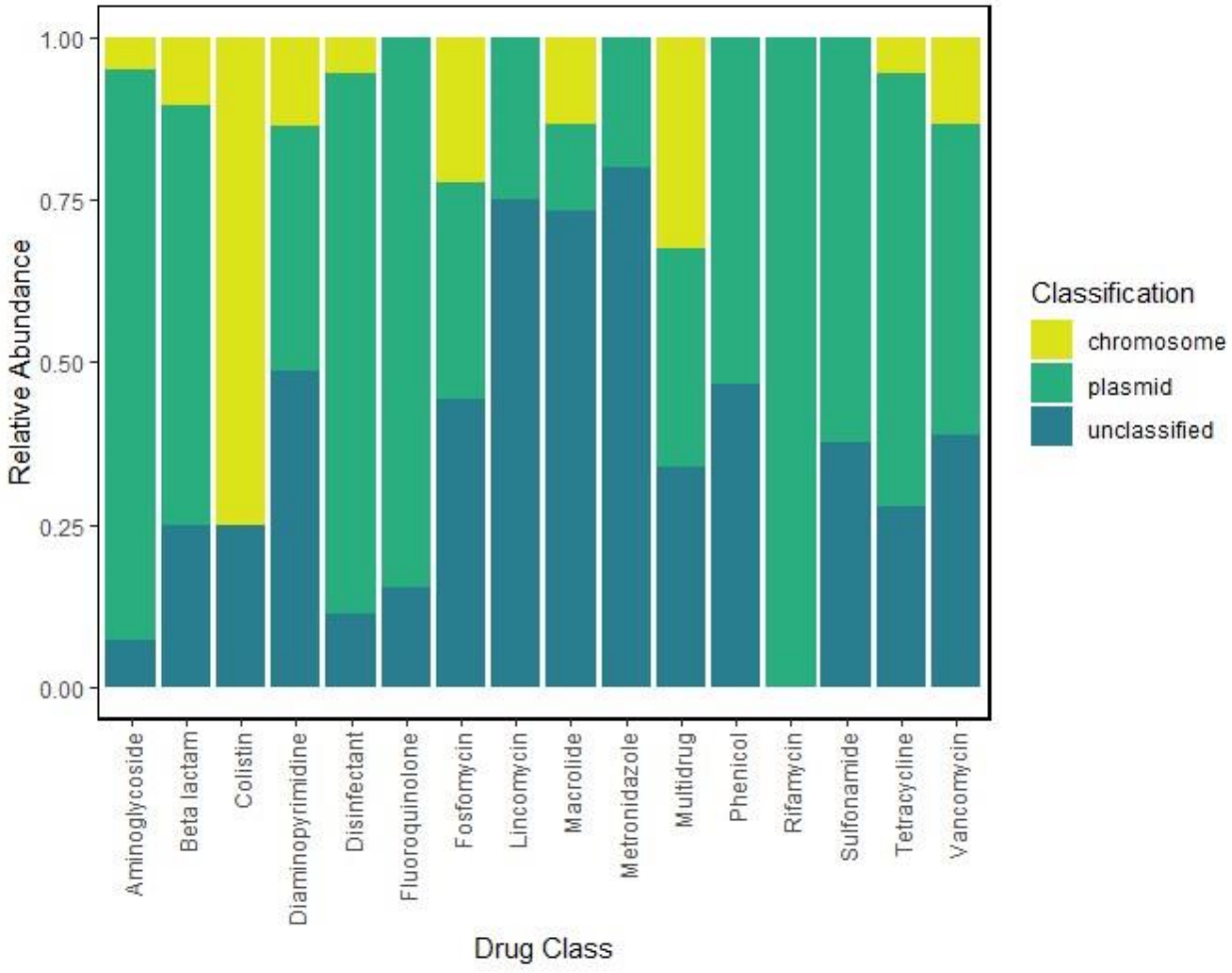


Figure 6. Relative Abundance of Chromosome and Plasmid-Associated Antimicrobial Resistance Genes: Chromosome and plasmid-associated ARGs were identified by PlasFlow, and classified into drug classes designated by CARD.

Of the ARG-like ORFs detected, we identified 66 chromosomal-associated and 212 plasmid-associated genes (Figure 7). The majority of drug classes were composed of plasmid-associated genes. All 7 genes associated with resistance to rifamycin were classified as plasmid. Identified rifamycin resistance genes were *arr-2, arr-3,* and *arr-8*, which are in plasmids of *P. aeruginosa, E. coli,* and other *Enterobacteriaceae*. Colistin resistance had the greatest abundance of chromosomal genes (75%). The top five ARGs identified across all samples were *vanW* gene in the *vanI* cluster (6.7%), *rsmA* (6.5%), *vanY* gene in the *vanB* cluster (5.3%), *adeF* (4.9%), and *ArnT* (3.1%). Genes *adeF* and *ArnT* were primarily associated with chromosomes; *vanW* in the *vanI* cluster, *vanY* in the *vanB* cluster, and *rsmA* were associated with plasmids (Figure 6, Figure 7).

Figure 7. Heatmap of Top 50 ARGs by Suburb: Hierarchical clustering of the relative abundances of the top 50 ARGs found in wastewater samples from each suburb in Harare.

2.4.3 Virulence factor genes

Out of 3580 curated genes within 1381 VFs in VFDB, we detected 412 genes within 373 VF proteins in the samples (Supplementary Figure 7). Increased Shannon-Wiener index scores were observed in low-income neighborhoods compared to high-income neighborhoods, but this difference was not significant (Supplementary Figure 4C, $p = 0.7$). Wastewater samples from high-income suburbs appear to cluster tightly together, while the low-income suburbs were very dissimilar from each other (Supplementary Figure 5C). There was no significant difference in the Bray-Curtis dissimilarity distances between high-income and low-income neighborhoods ($p = 0.6$).

A wide range of virulence factor genes was identified in the wastewater sample. Of the 14 VF classes, we detected 12. Classes Delivery, Adherence, and Motility were the most common, making up 33.3% (137), 20.4% (84), and 19.4% (80) of genes detected from all samples, respectively. Delivery, motility, adherence, and metabolic VFs were found in all suburbs (Figure 8). Classes exoenzyme, competitive advantage, regulation, and dispersion (<10 genes) were rarely detected. As observed in Supplementary Figure 7, the most common VFGs coded for were proteins building into Type III secretion systems (14.1%), polar flagella (13.3%), Type VI secretion systems (11.7%), Type IV pili (11.2%), and siderophores (10.9%). All other VFs were rarely detected (<5%).

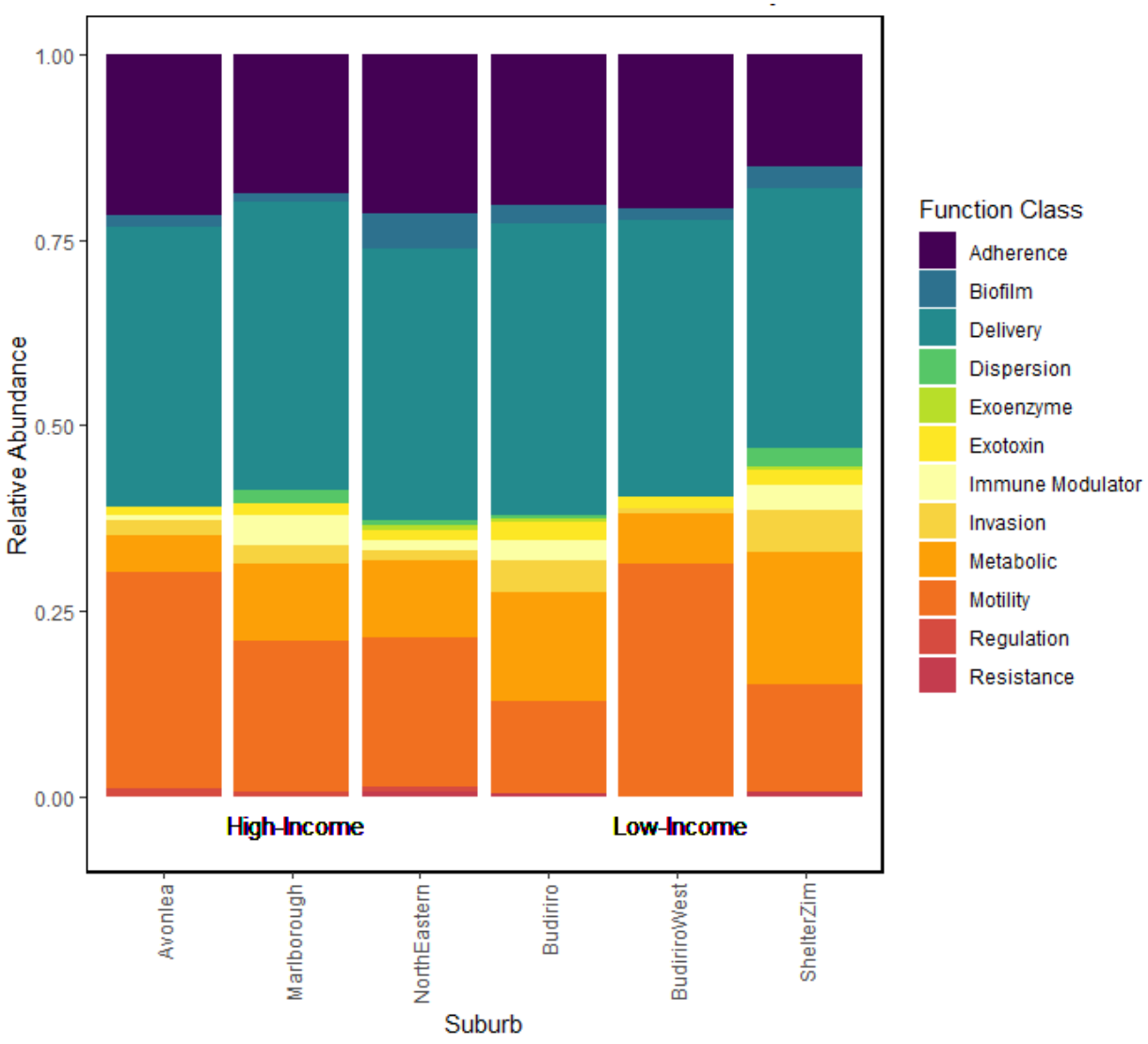


Figure 8. Relative abundance of Virulence Factor Function Classes: Virulence factor genes were classified into virulence function classes designated by the VFDB. Stacked bar chart displays the relative abundance of virulence factor function classes by suburb and socioeconomic status.

2.5 Discussion

2.5.1 Wastewater surveillance with shotgun metagenomics have potential as a population health monitoring tool for enteric diseases

In this pilot study, we described the enteric bacterial pathogens, AMR, and virulence factors in wastewater from six suburbs of Harare, Zimbabwe through shotgun metagenomics. Many previous studies demonstrated effective utilization of wastewater surveillance with PCR for monitoring poliovirus, typhoid, and SARS-CoV-2 (*Asghar et al., 2014; Kitakawa, Kitamura, Yoshida, 2023; Uzzell et al., 2023*). Few have employed metagenomic sequencing to characterize and monitor a multitude of infectious bacterial pathogens in wastewater. This study illustrates the potential of wastewater surveillance with shotgun metagenomics as a tool for population health monitoring of diarrheal diseases.

Alpha diversity of these wastewater samples was within the range of Shannon diversity scores (3.8 to 5) observed in previous literature on wastewater influent and sewage in sub-Saharan Africa (*Schneeberger et al., 2019; Jespersen et al., 2023*). The significant increase in the alpha diversity of wastewater from low-income suburbs compared to high-income suburbs could be explained by increased exposure to unsafe drinking water, which contains increased microbial diversity. Ingestion of poor-quality drinking water may lead to gut colonization and excretion of a larger diversity of microorganisms. Wastewater composite samples will reflect this if residents in the low-income suburbs are on the sewage system. Despite this significant difference, microbial and pathogen composition was similar and relatively stable between neighborhoods with differing socioeconomic status. These results suggest the presence of enteric pathogens may be a ubiquitous issue throughout Harare, regardless of socioeconomic status.

Major diarrheal pathogens such as *Salmonella sp.*, *Shigella sp.*, *Escherichia coli*, and *Vibrio cholera* were not the dominant species identified in these wastewater samples. Over 50% of the microbial composition can be attributed to *Aeromonas* and *Arcobacter sp.*, which are typically regarded as normal environmental bacteria found in freshwater and wastewater. However, these genera contain species known to carry virulence factors that can cause uncommon cases of mild to moderate gastroenteritis in humans, including *Aeromonas hydrophila* and *Arcobacter cryaerophilus* (*Ramees et al., 2017; Solís-Sánchez et al., 2023*). *A. cryaerophilus* has been gaining recognition as an emerging food-borne pathogen with growing antibacterial resistance, but large research gaps remain about its pathogenesis, distribution, and population impacts (*Ramees et al., 2017*). Although *Arcobacter sp.* is not known to cause severe, life-threatening diarrheal disease, they may still inflict public health impacts on communities with a high prevalence of malnutrition and HIV infection. Children under the age of 5 suffering from malnutrition and/or HIV infection are weakened and immunocompromised, leaving them highly susceptible to persistent or severe diarrheal disease, perpetuating malnutrition (*Ferdous et al., 2013; Pavlinac et al., 2015*). *Arcobacter* infections may contribute to this cycle, warranting the need for further research about this emerging pathogen and additional diarrheal pathogens previously considered inconsequential in public health.

Like the distribution of enteric pathogens, antimicrobial resistance may be a ubiquitous issue in Harare. ARGs revealed high multidrug resistance, in addition to beta-lactam, vancomycin, and tetracycline resistance in wastewater from all suburbs. Increasing drug resistance among *M. tuberculosis*, *E. coli*, *S. aureus*, and Gram-negative bacteria has been recognized as a growing issue in Zimbabwe, but high vancomycin resistance in these samples deviates from previous

literature (*Mauchaza et al., 2016; Mhondoro et al., 2019*). Between 2012 to 2017, increasing resistance to amoxicillin, ampicillin, and ciprofloxacin was attributed to these being the most prescribed antibiotics, while vancomycin resistance was of low concern (*Mauchaza et al., 2016; Mhondoro et al., 2019*). However, this study aligns with the antibiotic resistance displayed in the 2018 cholera outbreak. The *V. cholerae* strain responsible for the widespread outbreak revealed development of multidrug resistance, demonstrating intermediate to full resistance to ciprofloxacin, tetracyclines, beta lactams, 3rd generation cephalosporins, and sulfonamides (*Mashe et al., 2023*).

Vancomycin and last-line antibiotics have historically been considered expensive antibiotics in LMICs, leading to low use in Zimbabwe (*Mhondoro et al., 2019*), but the increased resistance observed in this study may be explained by certain practices and attitudes surrounding antibiotic use. Overprescription, poor patient adherence to antibiotics, inexpensive and easy access to substandard or falsified drugs, and the sale of antibiotics without prescription in private pharmacies are suspected to have contributed to rampant antibiotic resistance in Zimbabwe (*Gwatidzo et al., 2017; Olaru et al., 2022*).

This study also observed resistance in colistin, a polymyxin drug considered as a last-resort drug against multidrug-resistant Gram-negative bacteria. Colistin resistance is associated with widespread use in the animal agriculture industry to control disease and promote growth (*Sharma J. et al., 2022*). In Zimbabwe, antibiotic use for backyard poultry farms is not managed or regulated. Colistin-resistant *E. coli* has been isolated from poultry and colistin-resistant *S. enteritidis, S. epidermidis,* and *S. intermedius* strains were found in abattoir effluent (*Takawira et al., 2022; Gufe et al., 2021*). Effluent from backyard animal farms or local animal processing facilities that drain into the wastewater may explain the presence of ARGs against colistin. However, the majority of the ARGs coding resistance against colistin were chromosomal, indicating a low concern for the quick spread of colistin resistance among bacteria.

In contrast, 212 of the ARGs were associated with plasmid, suggesting many of the classified genes are carried on mobile genetic elements, causing concern for easy spread of antimicrobial resistance. Plasmid-associated ARGs coding for drug classes rifamycin, aminoglycoside, beta-lactam, fluoroquinolone, tetracycline, and sulfonamide are most concerning in playing a significant role in spreading resistance against these drug classes.

These ARG findings provide an initial update to past knowledge on AMR trends in Harare. Recognizing the rise in resistance to vancomycin and colistin can refine or expand the detection catalogue of current AMR surveillance efforts, such as the Zimbabwe One Health surveillance system. Findings can also guide new studies evaluating expression of ARGs against specific drug classes in clinical isolates, validating these preliminary results.

This study found a wide range of virulence factors and similar diversity between high-income and low-income suburbs. Type II, III, and VI secretion systems composed most of the delivery proteins (94%). All three secretion systems are key virulence factors utilized by many Gram-negative bacteria for an impressive range of functions, aiding disease progression and transmission (*Green & Mescas, 2016*). Type II secretion systems (T2SS) secrete additional virulence factors which are essential for environmental survival, such as proteases and metabolic enzymes for nutrition acquisition, in addition to disruption of host cells, such as aerolysin and cytolysin toxins (*Korotkov and Sandkvist, 2019*). *Salmonella* and *Shigella sp.* require Type III secretion systems (T3SS) to invade and establish in non-immune host cells, while EPEC and

EHEC may use T3SS to interfere with immune response and to attach to epithelial cells (*Deng et al., 2017*). Type VI secretion systems (T6SS) have been detected in *Shigella sonnei*, *Pseudomonas aeruginosa*, and *Vibrio cholerae*, with the primary function of interbacterial competition to facilitate host colonization and environmental survival (*Coulthurst, 2019*). *Shigella spp.* also utilize Type VI secretion systems to inject toxins into target host cells.

Detection of certain VFGs supports the potential presence of pathogenic strains, such as *Salmonella enterica serovar typhi* and pathogenic *E. coli*. For example, the Virulence capsular polysaccharide (Vi) antigen is a key virulence factor of *Salmonella enterica serovar typhi*, aiding in evasion from the immune system (*Parween, Yadav, Qadri, 2019*). This protein was detected in five of the six neighborhoods, and *Salmonella enterica* was detected in all six neighborhoods. Detection of both the Vi antigen and the bacterium in the same suburbs can support the possible identification of *S. enterica serovar typhi*. Another example is pathogenic *E. coli,* which produce virulence factors that aid in colonization and host cell disruption that differ from non-pathogenic *E. coli* (*Kaper, Nataro, Mobley, 2004*). Toxin and autotransporter Pet, a protein of the subfamily serine protease autotransporters of enterobacteriaceae, was detected in one suburb that contained *E. coli* in the wastewater. This protein is produced by pathogenic *E. coli* and *Shigella* (*Kaper, Nataro, Mobley, 2004*); thus, the detection of both Pet and *E. coli* in wastewater may suggest the possible presence of pathogenic *E. coli* in the neighborhood.

The high abundance of VFGs coding for siderophores compared to other metabolic virulence factors (83%) may be attributed to heavy metal pollution in wastewater, surface-impacted waters, and groundwater in Harare (*Moyo, 2013*). Many pathogenic *E. coli* produce siderophores to accumulate iron, which is essential for metabolism and protein formation that aid in their survival and proliferation. Siderophore production has also been found to be correlated with ARG expression, particularly resistance to fluroquinolone and sulfonamide drugs (*Khazaal et al., 2022*). Pyoverdine, a siderophore specific to multi-drug resistant *P. aeruginosa*, can chelate iron from beta-lactam drugs such as cefiderocol, reducing drug efficacy (*Khan, Palmer, & Dillon et al., 2024*). In other Gram-negative bacteria, increased iron accumulation from siderophores can facilitate the activity of metallo-beta-lactamases against beta-lactam drugs (*Cahill et al., 2016*). High multidrug and beta-lactam resistance, coupled with the high abundance of siderophore-related genes, observed in Harare could be attributed to both metal and antibiotic pollution of the waters.

In contrast, low detection of VFGs coding for invasion, immune modulation, and exotoxin proteins was observed, suggesting a low number of potential pathogens harboring these proteins. However, a single detection of Shigella enterotoxin ShET2 can be significant, as this indicates the presence of *Shigella sp*. in the community wastewater. A small load of 1 to 10 cells of *Shigella sp.* can cause diarrheal disease in humans.

The diversity of ARGs and VFGs within suburbs of the same income status were both higher compared to the microbial diversity within suburbs of matching income status. This finding is similar to previous literature examining the wastewater microbiome, resistome, and present virulence factors. Studies investigating the wastewater resistome identified a range of Shannon's diversity scores between 3.3 and 4.2 within wastewater (*Lee et al., 2023*). A previous study conducted in China observed a high alpha diversity range of ARGs in wastewater from urban wastewater treatment plants, between 6.55 to 7.22 for Shannon's diversity scores (*Zhang et al., 2016*). The range of Shannon diversity scores observed for VFGs in wastewater from this study

is lower than the findings in the previous study (4.5 to 5.2). However, this is not unusual, as our samples were collected upstream of wastewater treatment plants, and Harare is a smaller city compared to the sampled cities in China. ARGs and VFGs provided additional value to pathogen identification, highlighting microorganisms of concern through their potential to spread ARGs and their ability to cause moderate or severe disease with certain VFGs.

This study contains a large breadth of information regarding pathogens, AMR, and virulence factors. The compilation of this new background data can serve as the foundation for future exploratory studies, targeted interventions, and routine wastewater surveillance with shotgun metagenomics.

2.5.2 Limitations

This pilot study has the following limitations, many of which can be attributed to its budget constraints. At the time of sample collection, the study was only able to collect and analyze three samples from six suburbs for a total of eighteen samples. This is a small sample size from a limited selection of suburbs. This sampling bias can explain large variabilities observed in the wastewater microbial composition, ARGs, and VFGs between samples, which may affect identification of significant differences. However, this study remains significant, as the primary study aim, to identify enteric pathogens and to characterize relevant ARGs and VFGs, was achieved. This information serves as the foundation for future research.

The samples were collected at a few timepoints in one season, introducing temporal bias into this study. The lack of a longitudinal analysis or a comparison to another season does not allow for identification of differences in pathogen patterns, leading to missed information on how seasonality may affect pathogen characteristics and disease burden. However, results provide a reference of the dry season for a future study characterizing pathogens, ARGs, or virulence factors in wastewater over multiple seasons.

Another concern was the lack of reported clinical cases of diarrheal illness with identified pathogens as reference to assess our findings. In general, there was little information to provide context to our results. The study indicates the presence of enteric pathogens and ARGs in the Harare neighborhoods, but real-time public health importance cannot be determined without confirmed cases.

Specific results on pathogen and ARG compositions and patterns may only be applicable to Harare, Zimbabwe, as the environmental microbiome in wastewater may vary between cities and LMICs due to different WASH infrastructure, health policies, culture, geography, and many additional environmental and behavioral factors. However, the strong ability findings that evaluate the performance of the BMFS recovery with metagenomic analysis as strong tools for environmental surveillance can be relevant in many communities.

Batch effect analysis issues are also a concern, due to multiple sampling events over several weeks that differ between high-income suburbs and low-income suburbs. This limitation was addressed with the selected diversity indices and statistical tests. In addition, batch sampling is commonly conducted in environmental microbiology; thus, confounding will have been considered as part of the sampling design and approach. An analysis of this detail has not previously been conducted in Harare suburbs. Therefore, the information from this study to characterize diarrheal disease burden in this community would be beneficial.

Lastly, unlike whole genome sequencing, which requires culturable isolates, metagenomic sequencing cannot discriminate DNA from live or non-viable organisms. In addition, without culture, phenotypic expression of ARGs and VFGs cannot be confirmed. Sequencing can only detect the presence of a gene, but not its expression. Another limitation of metagenomic sequencing analysis is reduced resolution with lower-level taxonomic classification, which may prevent distinguishing pathogenic strains and environmental microorganisms *(Shen et.al, 2021*). The most reliable classification can be made at the genera level, but this information is inadequate for pathogen identification. This concern was overcome by comprehensive analysis linking identified pathogens with VFGs, as the presence of specific species and pathogenic strains can be suggested by detected virulent characteristics.

2.5.3 Conclusion Statement

From this study, we concluded that wastewater surveillance paired with shotgun metagenomics has the potential as a public health monitoring tool that can supplement clinical surveillance. This tool has also shown the possibility to simultaneously monitor enteric diseases and AMR in Zimbabwe urban communities. The tool did not demonstrate the ability to detect significant differences in pathogen composition, AMR, and virulence factors of wastewater from urban neighborhoods of differing socioeconomic status.

2.5.4 Next Steps

Validating these findings and exploring new questions from this study will further improve our understanding about enteric disease burden, guiding public health guidelines and interventions. Future studies beyond this pilot project can take several approaches. A cross-sectional study with a larger sample size and more sampled neighborhoods will have a strong focus on investigating spatial and socioeconomic differences in exposure. In comparison, a longitudinal study with clinical reports for reference will aim to further demonstrate that wastewater surveillance with shotgun metagenomics can be an effective tool for population health monitoring of enteric diseases. For both aforementioned study designs, a large sample size is necessary to validate any calculated significance in identified patterns, to improve confidence in attributing variation to a factor, and to evaluate the ability to process large sample sizes in a timely manner that ensures results still hold public health importance. Other studies may investigate detection of specific pathogens and antimicrobial resistance genes relevant to the community, utilizing simple, inexpensive, targeted, and rapid detection methods that can be quickly and easily implemented in routine public health surveillance.

Exploratory studies could examine tentative findings that arose from this study. For example, with the possible concern of animal exposure through wastewater discharge into water bodies, future research can be structured through a One Health lens to improve understanding of the roles animals and humans play in water safety and their health. The Zimbabwe One Health surveillance system could consider a study investigating colistin resistance in wastewater discharge and chickens from poultry operations. Another study may explore resident WASH and health behaviors in neighborhoods of differing socioeconomic status to understand risk behaviors that may contribute to differences in pathogen and AMR patterns observed through wastewater.

Feasibility and practicality of implementing a routine wastewater surveillance program with shotgun metagenomic sequencing depends on funding, relevance, and resources. Wastewater provides simple composite samples of the community, while sequencing can yield a large

amount of data from wastewater samples. Establishing this tool in a high-income country with abundant funding and resources for a high-complexity laboratory can be feasible. Finding trained laboratory scientists and bioinformaticians may be the primary concern. In a low-income country, there may be insurmountable barriers preventing the addition of this method in routine public health surveillance. There are very few to no laboratories with the capacity and capability to conduct next-generation sequencing. Sending out samples for sequencing can be expensive. Tight budgets may impede acquisition of funding for an expensive and resource-intensive public health monitoring tool. Most importantly, wastewater surveillance with shotgun metagenomic sequencing is most practical for monitoring multiple enteric pathogens of high clinical and AMR concern in the general community.

FUNDING
This study was supported by the UW DEOHS Microbiome Health Initiative.

AUTHOR CONTRIBUTIONS
V.R., S.M., and N.Z. conceptualized the study, S.M., N.Z., and K.L. reviewed and edited the article. V.R., I.N., J.M., B.M., G.N., P.C., A.M. performed sample collection and fieldwork. N.Z. and A.O. performed sequencing. S.Y.C. performed data analysis and result interpretation, and wrote, reviewed, and edited the article.

DATA AVAILABILITY STATEMENT
Data is available upon request.

CONFLICT OF INTEREST
The authors declare there is no conflict.

CHAPTER 3: CONCLUSIONS

3.1 Conclusion Statement

From this study, we conclude that wastewater surveillance paired with shotgun metagenomics has the potential as a public health monitoring tool that can supplement clinical surveillance. This tool has also shown the possibility to simultaneously monitor enteric diseases and AMR in Zimbabwe urban communities. When considering differences in socioeconomic status, the tool was not able to detect significant differences in pathogen composition, AMR, and virulence factors.

3.2 Reflections

The current findings from this study have been a small achievement towards realizing shotgun metagenomics in routine wastewater surveillance for enteric disease monitoring in communities. Very few studies have simultaneously examined enteric disease and AMR burden in urban wastewater, and none have previously conducted this type of study in Harare. Although the author was not involved in sample collection and processing, both the data collection and data analysis phases required great effort to determine the best approaches for conducting wastewater surveillance with molecular techniques in a feasible manner.

Further in-depth analysis with new programs or approaches can still be conducted on the sequences to expand our results. However, there are three major challenges that were and may continue to be encountered throughout the bioinformatics, data analysis, and interpretation: 1) output from one database or program is not easy to merge with output from another application, 2) lack of a standard approach for shotgun metagenomic analysis, and 3) database limitations. 1) Linking results from Kraken2, CARD, and VFDB is not possible by contig, as the output from each of these databases do not share similar identifiers to easily merge outputs together. The best alternative method was to connect the results by sample and suburb. Missing the linkage between More granular results are lost2) There are a variety of different approaches for bioinformatic analysis of shotgun metagenomic sequences, from pre-processing to statistical testing. To improve read depth and genome recovery, reads from several samples may be co-assembled prior to downstream analysis (*Rajeev et al., 2023; Ma et al., 2023*). A few steps are uniform across most current literature, such as the selection of alpha and beta diversity indices (Shannon's diversity index and Bray-Curtis dissimilarity index). Many software packages can perform a few bioinformatics steps or format outputs compatible for a few select packages to perform the next step in your analysis. In addition, packages are not written in one coding language, requiring competency in multiple languages. 3) Databases are limited based on submitted sequences, curation criteria, and maintenance. For example, the VFDB is a highly relevant database for virulence factors as it is updated weekly, and sequence submission is greatly encouraged. However, the database is limited to sequences to well-studied virulence factors from clinically relevant pathogens, which is not comprehensive enough for sequences from environmental samples. CARD also contains this issue, as AMR is tied with clinical relevance, but the RGI overcomes this problem with its criteria for strict hits, identifying contigs similar to the reference sequences as potentially functional ARGs (*Hyatt et al., 2010; Alcock et al., 2023*). Another example is Plasflow. Although this database has been reliable for classification of submitted contigs as plasmid or chromosomal, it has not been updated since 2018. Each database will have its strengths and weaknesses. Selection of the most ideal database depends on the sample types, sequencing target, and aims of the study.

Regarding actionability in public health, these findings are promising in demonstrating the effectiveness of this public health monitoring tool, but there are many barriers that may prevent its implementation, particularly in LMICs. Feasibility and practicality depend on funding, relevance, and resources. Establishing this tool in a high-income country with abundant funding and resources for a high-complexity laboratory can be achievable within a few years. During COVID-19, the University of Oklahoma established a wastewater surveillance program and tracked norovirus, *Salmonella spp.*, and *Campylobacter spp.* for 18 months with Oklahoma State Department of Health weekly reported cases as reference (*Kuhn et al., 2023*). This study is a proof-of-concept for wastewater surveillance and molecular detection of pathogens at the state-level. In a low-income country, there may be a number of barriers preventing the addition of this method in routine public health surveillance for one to two decades. Next-generation sequencing costs may have decreased, but few local labs in Zimbabwe have the capability and capacity to conduct this method. Sending out samples to third-party laboratories will increase turnaround time, losing the aspect of near real-time information about pathogen distribution and disease burden in the community. Additional funding is also necessary to run a new operation for a method that remains both expensive and resource-intensive in LMICs. In both high-income and low-income countries, expertise and training are important to consider in calculations for funding and resources. Trained laboratory professionals in sequencing, a bioinformatician, and sufficient computing power must be considered when implementing this new method in a public health laboratory. Interpretation of sequencing results requires bioinformatics expertise and ample computing power. Lastly, wastewater surveillance with shotgun metagenomic sequencing is practical for monitoring both multiple enteric pathogens of high clinical concern and relevant ARGs in the general community.

3.3 Significance

This pilot study provides value as a foundational study for several different research approaches that share the goal to improve diarrheal disease burden in the most affected communities. The findings demonstrated wastewater surveillance paired with shotgun metagenomics can identify enteric pathogens and relevant genes from the mass of information extracted from the samples. In general, this study supports the potential of these two methods as a public health monitoring tool in many communities, contributing to the growing body of literature on wastewater surveillance for multiple pathogens. For Harare, identified enteric pathogens and ARGs provide a starting point for refining targets of current surveillance systems. Detection of key virulence factors identified from this study can also assist with pathogen identification in targeted surveillance.

Pursuing further research stemming from this study may lead to long-term outcomes that improve diarrheal disease burden. Development of simple, inexpensive, rapid, and targeted tests for pathogens of high clinical and AMR concern in wastewater can provide near real-time information that guides public health interventions and clinical treatment in specific communities Outbreaks may be more quickly controlled or prevented with the implementation of targeted testing. Routine wastewater surveillance with sequencing may provide updates about emerging trends with pathogens and AMR to public health officials, acting as baseline data that may trigger additional action by public health officials. Growing the body of data in previously under-researched countries is essential to present relevant information that encourages action from policymakers to develop policies that reduce AMR and diarrheal disease burden.

ACKNOWLEDGMENTS

I would like to acknowledge the Environmental and Occupational Health Microbiology Laboratory for their unwavering support and helpful guidance throughout the data analysis of this project. Thank you to Dr. Vurayai Ruhanya and collaborators at the Department of Medical Microbiology of the University of Zimbabwe, and the City of Harare. I would not have obtained the data from the Virulence Factor Database without the help of Dr. Kelsey Jesser. I also acknowledge the help from the DEOHS IT team, John Yocum and Brian High, for assisting me by troubleshooting errors with the bioinformatics software and packages. Lastly, thank you to the UW DEOHS Environmental Health Microbiome Initiative for the initial funding of this project.

Supplementary Information

SI 1. Detailed Methods

SI 1.1 Bag-Mediated Filtration System

Due to the sheer volume of water and contaminants from multiple sources, the concentration of microorganisms is highly diluted in wastewater and wastewater-impacted surface waters. Sampling methods that can retain and concentrate the present microorganisms can overcome this primary drawback of wastewater sampling. The bag-mediated filtration system (BMFS) is a novel sampling method originally designed by the Environmental and Occupational Health Microbiology Laboratory at the University of Washington to conduct environmental surveillance for poliovirus (*Fagnant et al, 2014*). This method includes a specially designed elution device made to accommodate custom Virocap filter housings (*Fagnant et al, 2017*). A study in Pakistan demonstrated that the original version of the BMFS improved poliovirus detection compared to the two-phase separation method (*Zhou et al, 2018)*. Another study in Kenya validated refinements to a second version of the BMFS, demonstrating that a double elution combined with a skimmed milk flocculation protocol gave improved recoveries (*Fagnant et al, 2018*). The second edition of the BMFS was used to conduct environmental surveillance for enteroviruses and poliovirus in Harare, Zimbabwe (*Ruhanya et al, 2022*).

SI 1.2 Sample Collection

Samples were filtered through the BMFS using gravity by hanging sampling bags on a tripod, then the remaining water was dispensed back into the water source. Once the water had run through the filter and the filter housing is empty of liquid, the housing was disconnected from the field tubing and wiped down with bleach, end caps attached to inlet and outlet ports, and sample placed inside secondary containment in a cooler on ice until reaching the University of Zimbabwe Virology Laboratory.

SI 1.3 Elution

A solution of 1.5% beef extract and 0.05M glycine (150ml, pH 9.5) was injected into the inlet of the housing, then held for 15 minutes at room temperature. After the allotted time, the bilge pump was used to aspirate the eluted sample into a collection cup. The injection step was repeated to perform a second elution. After the second waiting period, both collected volumes were combined into af larger plastic bottle, totaling around 300ml per filter housing sample. Collection containers were labeled with sample information, the pH adjusted to 7.0 -- 7.5 using NaOH and HCl, and stored in a refrigerator for secondary concentration within 24 hours.

SI 1.4 Secondary Concentration

Two millilitres of a 5% skim milk solution were added to each sample, adjusting the pH to between 3.0-4.0. The samples were then shaken at 200rpm for two hours at room temperature. Samples were separated into 50ml conical tubes and centrifuged at 3500G for 30minutes at 4^{o}C. The supernatant was discarded, and pellets were resuspended in 10ml of cold 1xPBS.

SI 1.5 Relative Abundance Plots for Microbial Composition

Relative abundance plots for microbial composition were composed from modifying a public R script for a Shiny app that visualizes merged Bracken data into relative abundance plots (bracken_plot).

SI 2. Tables and Visuals

Supplementary Table 1. Descriptive Information of the Wastewater Samples: States sample ID, suburb name, suburb density, suburb income status, sampling date, sampled water volume, outdoor temperature, and weather.

Sample	Suburb	Density	Income	Date	Volume (litres)	Temperature (°C)	Weather
S2	Northeastern	Low	High	4/19/2019	4.75	28	Sunny
S3	Northeastern	Low	High	4/27/2019	4.5	29	Sunny
S4	Northeastern	Low	High	5/3/2019	4.8	20	Cloudy
S5	Avonlea	Low	High	4/19/2019	3	22	Cloudy
S6	Avonlea	Low	High	4/27/2019	3.75	28	Sunny
S7	Avonlea	Low	High	5/3/2019	3.25	29	Sunny
S8	Marlborough	Low	High	4/19/2019	5.3	20	Cloudy
S9	Marlborough	Low	High	4/27/2019	3	28	Sunny
S10	Marlborough	Low	High	5/3/2019	3	29	Sunny
S11	Budiriro West	High	Low	4/23/2019	3.5	26	Sunny
S12	Budiriro West	High	Low	5/4/2019	5.5	30	Sunny
S13	Budiriro West	High	Low	5/9/2019	5.4	30	Sunny
S14	Budiriro	High	Low	4/22/2019	1.75	26	Sunny
S15	Budiriro	High	Low	5/4/2019	3.5	30	Sunny
S16	Budiriro	High	Low	5/9/2019	3	30	Sunny
S17	Shelter Zim	High	Low	4/22/2019	3	26	Sunny
S18	Shelter Zim	High	Low	5/4/2019	3	30	Sunny
S19	Shelter Zim	High	Low	5/9/2019	3.25	30	Sunny

Supplementary Table 2. Recorded Average Qubit Concentrations of DNA Extracts (ng/uL)

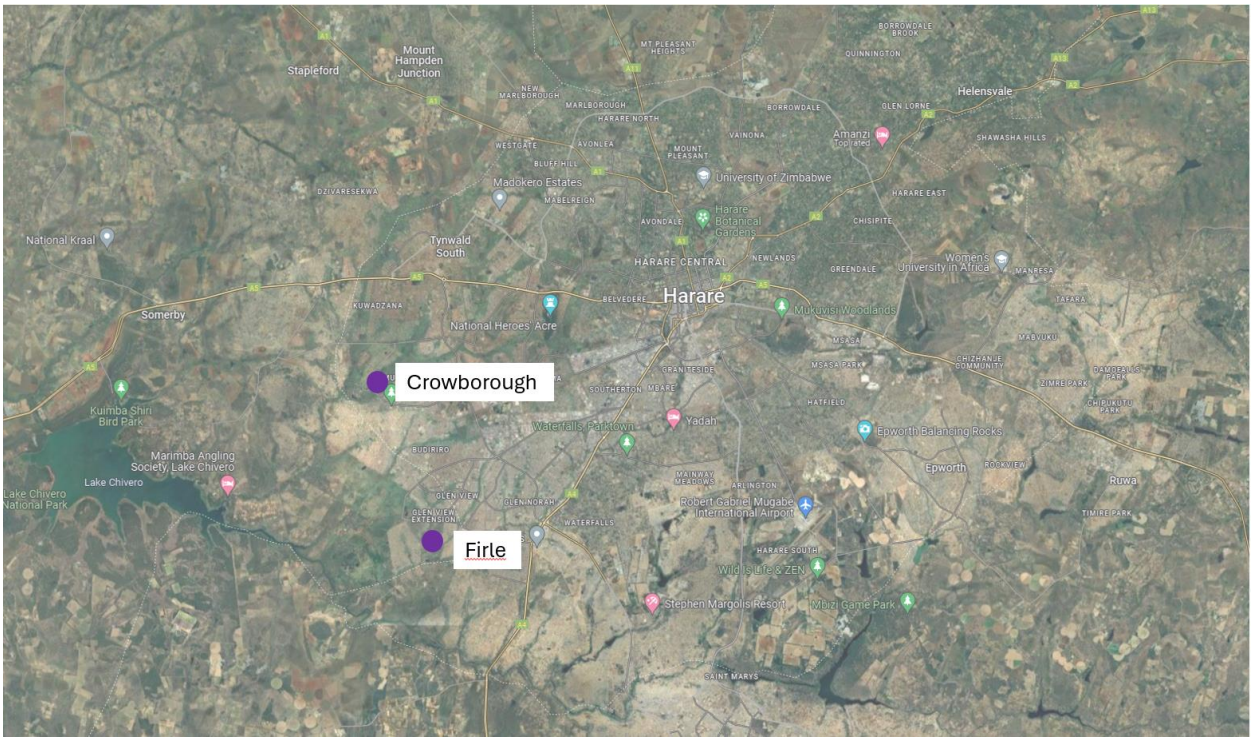
Sample	Average DNA Concentration (ng/uL)
S2	6.08
S3	39.20
S4	10.70
S5	27.03
S6	14.60
S7	0.86
S8	9.39
S9	10.30
S10	3.90
S11	38.73
S12	0.90
S13	2.57
S14	29.03
S15	5.10
S16	3.15
S17	28.90
S18	1.46
S19	7.83

Supplementary Table 3. Sequencing Quality Control/Quality Assurance

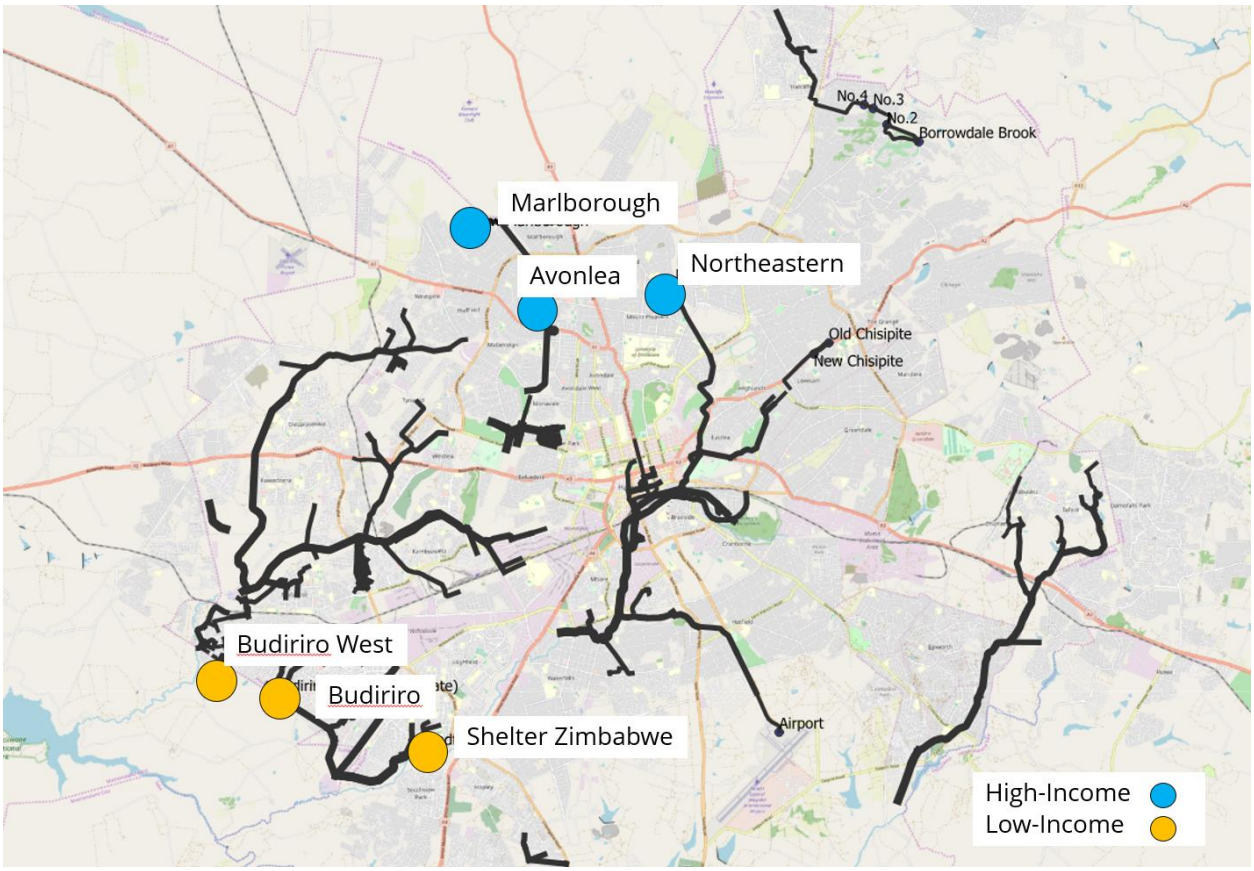
Suburb	Total Raw Read Count	Total Read Count (Trimmed)	Total Contig Count (metaSPAdes)
Avonlea	29,526,875	27,586,150	2,429,501
Marlborough	26,980,900	25,691,498	1,876,982
Northeastern	25,852,043	23,952,499	2,038,288
Budiriro	24,036,893	22,901,972	2,449,229
Budiriro West	14,419,790	13,923,729	1,119,668
Shelter Zimbabwe	30,540,056	29,170,050	2,058,541
Total	151,356,557	143,225,898	11,972,209

Supplementary Table 4: Illumina Adapter Sequences

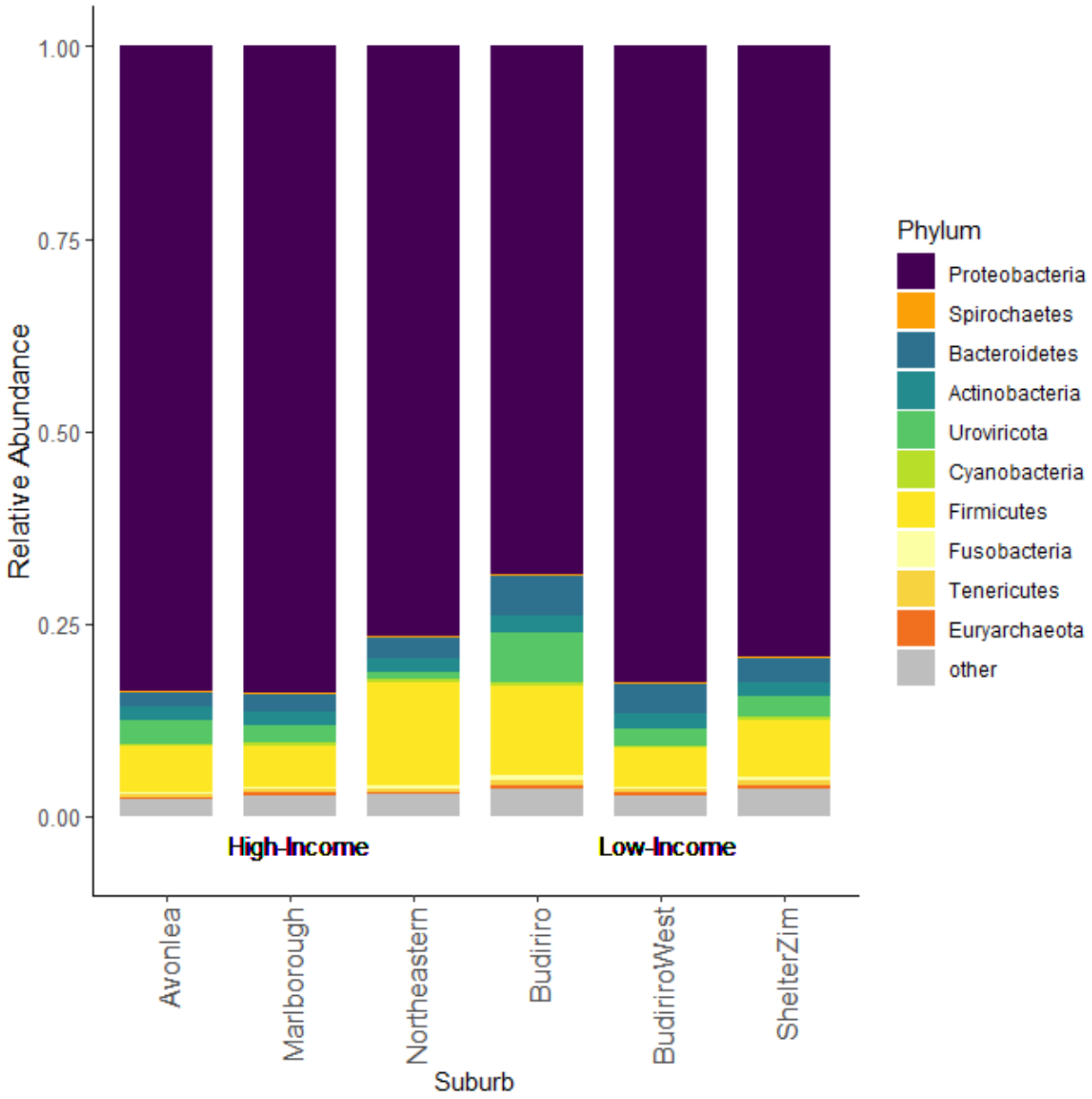
Sample ID	Sample Number	Index 1 (i7) Nucleotide Sequence	Index 2 (i5) Nucleotide Sequence
ZimEV_01	S2	CGGTTACGGC	CTATAGTCTT
ZimEV_02	S3	GAGAATGGTT	TTGCTGCCGA
ZimEV_03	S4	AGAGGCAACC	CCATCATTAG
ZimEV_04	S5	CCATCATTAG	AGAGGCAACC
ZimEV_05	S6	GATAGGCCGA	GCCATGTGCG
ZimEV_06	S7	ATGGTTGACT	AGGACAGGCC
ZimEV_07	S8	TATTGCGCTC	CCTAACACAG
ZimEV_08	S9	ACGCCTTGTT	ACGTTCCTTA
ZimEV_09	S10	TTCTACATAC	TTACAGTTAG
ZimEV_10	S11	AACCATAGAA	CCATCTCGCC
ZimEV_11	S12	GGTTGCGAGG	TTGCTCTATT
ZimEV_12	S13	TAAGCATCCA	AATGGATTGA
ZimEV_13	S14	ACCACGACAT	CCGCATACGA
ZimEV_14	S15	GCCGCACTCT	CGAGGTCGGA
ZimEV_15	S16	CCACCAGGCA	ATTCCATAAG
ZimEV_16	S17	GTGACACGCA	GTCCGTAAGC
ZimEV_17	S18	ACAGTGTATG	CCGTATGTTC
ZimEV_18	S19	TGATTATACG	TGTAATCGAC



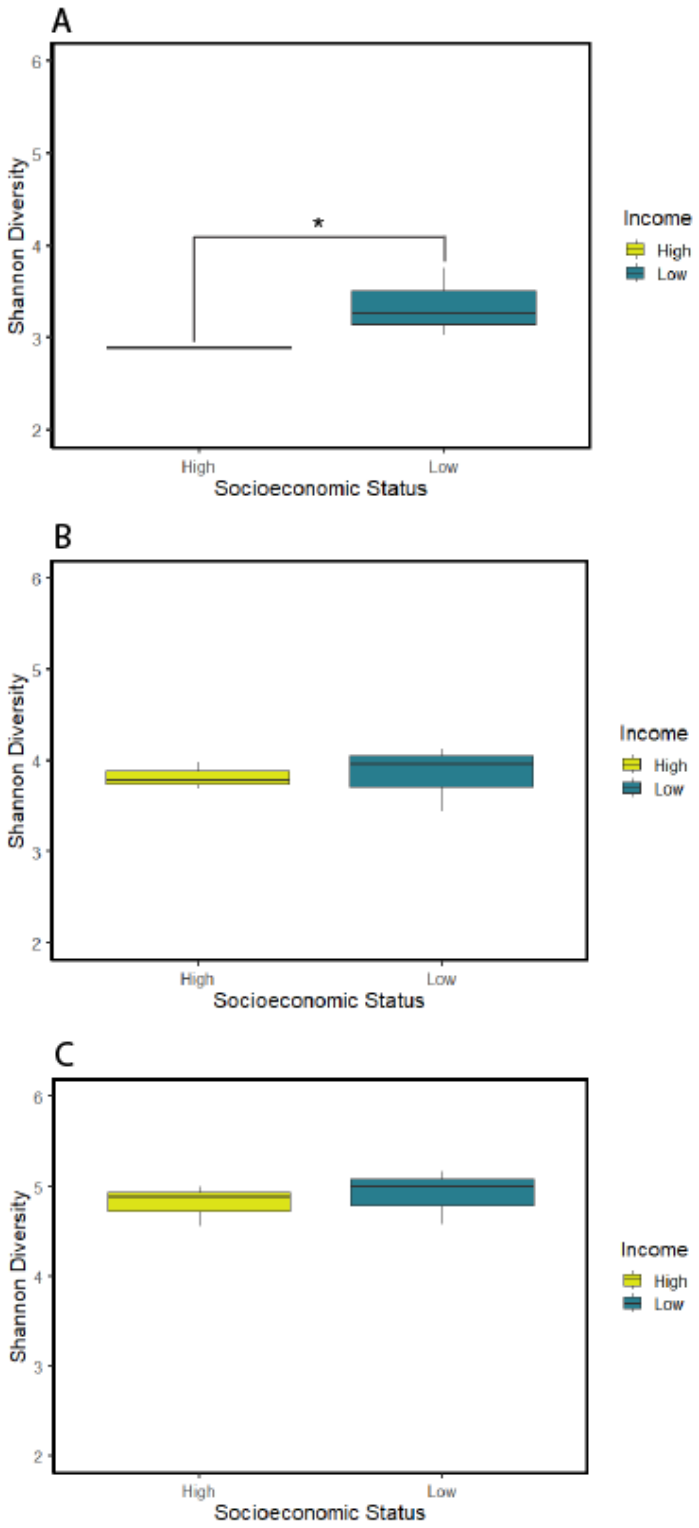
Supplementary Figure 1. Locations of two major wastewater treatment plants serving Harare, Zimbabwe. Figure was generated with Google Maps and Microsoft PowerPoint



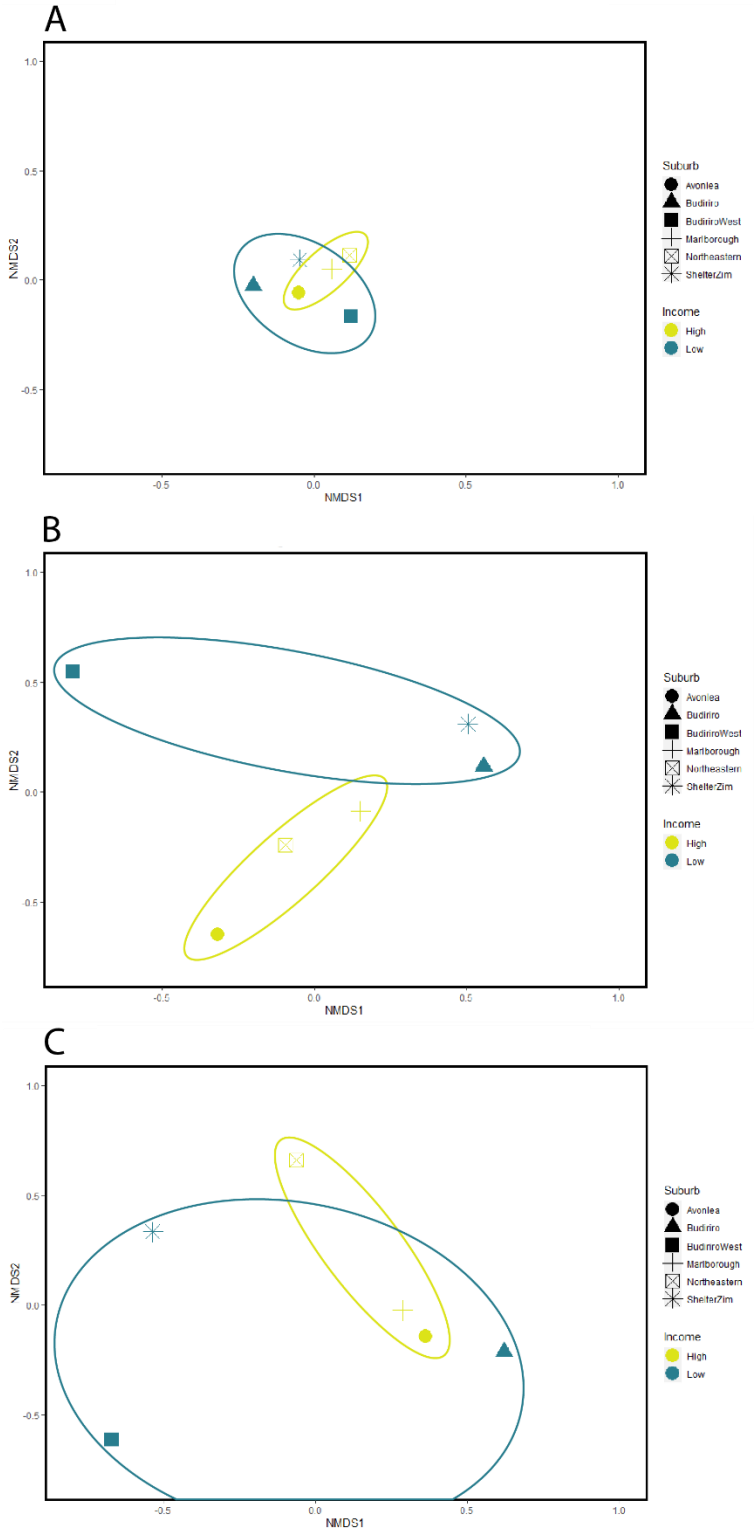
Supplementary Figure 2. Map of sampling locations: Sampling locations within each suburb are highlighted on the map. Dark blue lines indicate the route of wastewater pipes. Thickness of the lines indicate volume of wastewater.



Supplementary Figure 3. Relative Abundance of Top 10 Phyla: Relative abundance of the top 10 bacterial and archaeal phyla in wastewater samples by suburb and income. Phyla not included in the top 10 were grouped into "Other".

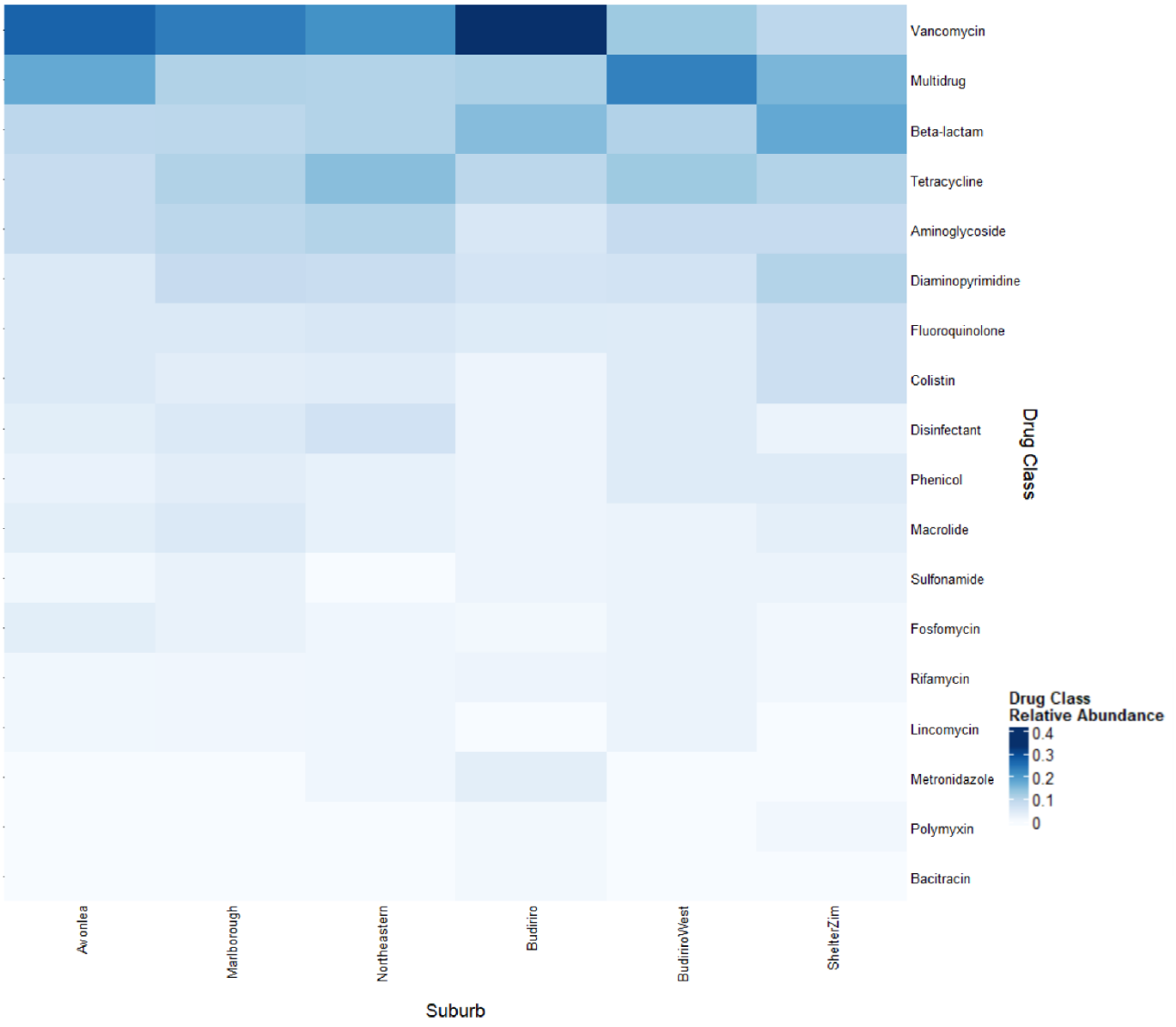


Supplementary Figure 4. Comparison of Shannon Diversity Scores between High-income and Low-income Suburbs: A) Metagenomes B) Antibiotic Resistance Genes C) Virulence Factor Genes
Samples were merged by respective suburb prior to data processing and bioinformatics.

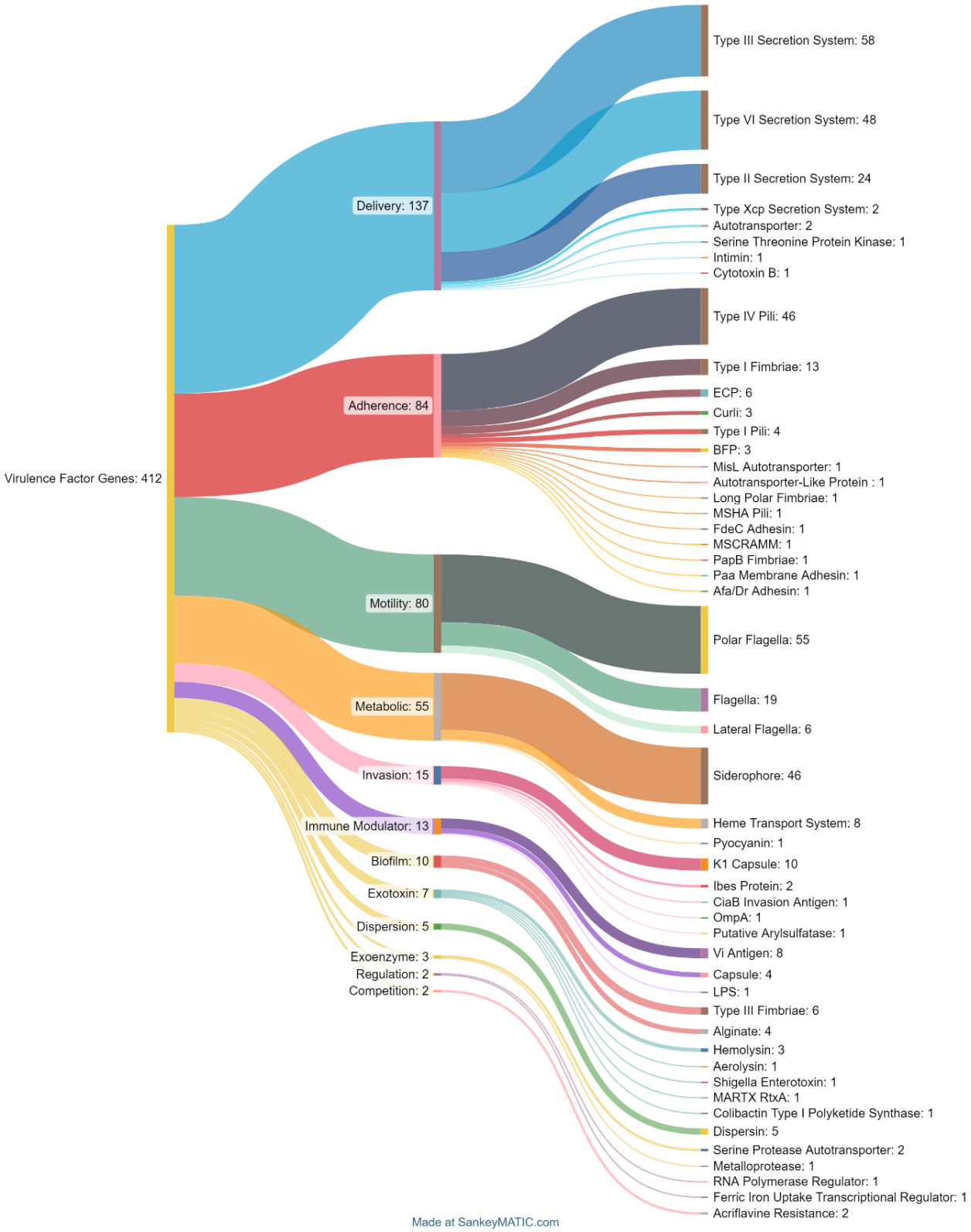


Supplementary Figure 5. Comparison of Bray-Curtis Dissimilarity Distances between High-income and Low-income Suburbs: A) Metagenomes B) Antibiotic Resistance Genes C) Virulence Factor Genes

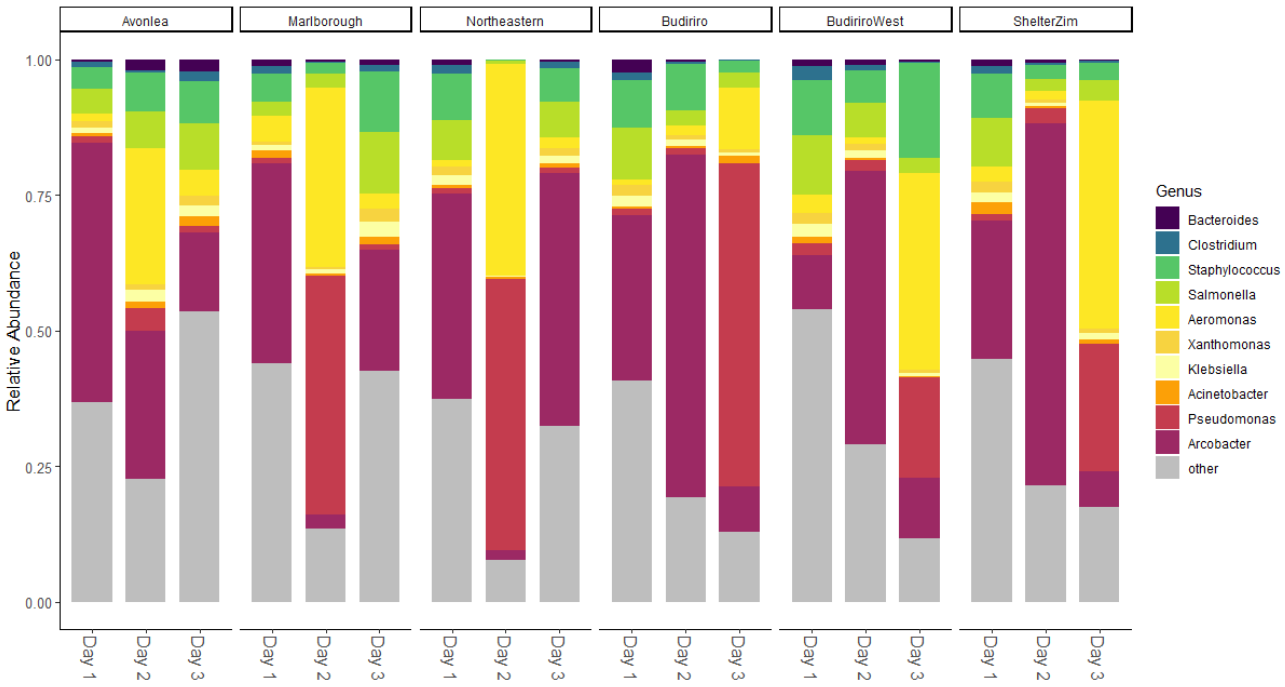
Samples were merged by respective suburb prior to data processing and bioinformatics.



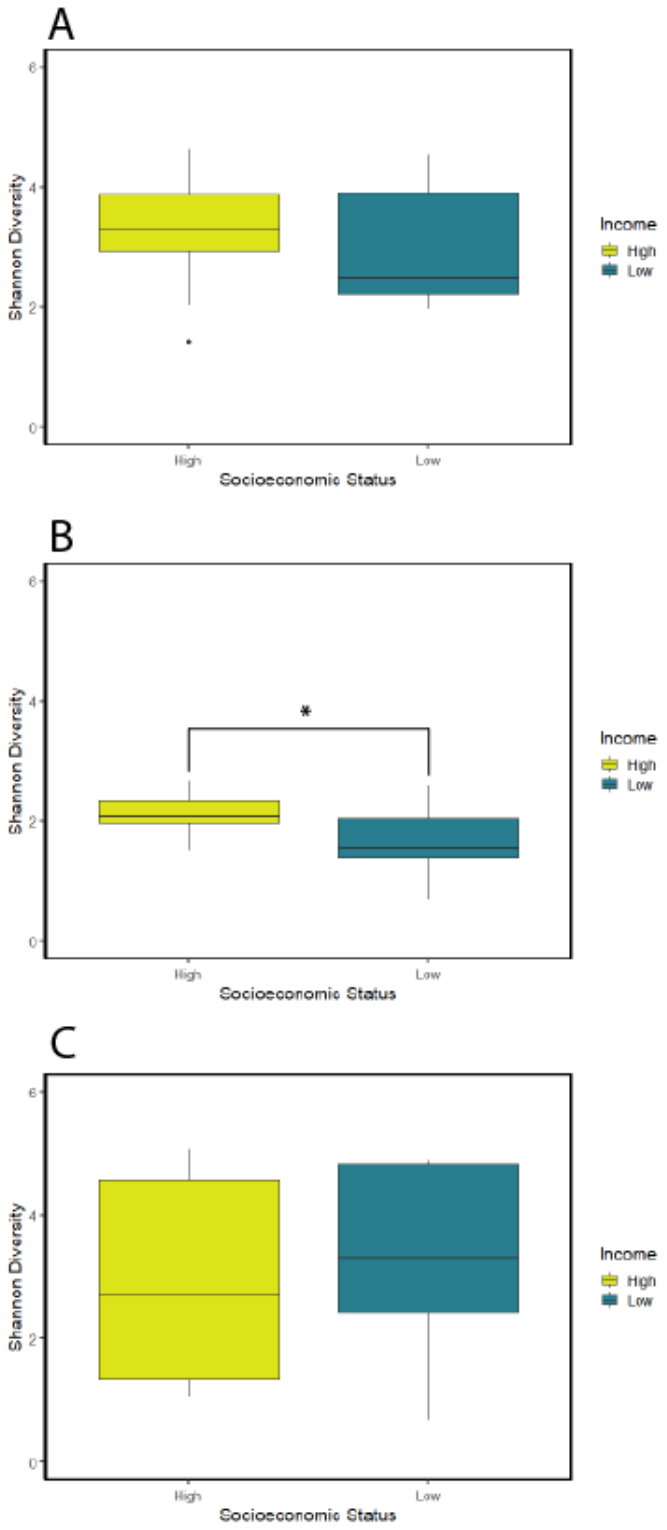
Supplementary Figure 6. Heatmap of Drug Class Resistance by Suburb: Hierarchical clustering of ARG drug class relative abundance by suburb, visualized as a heatmap.



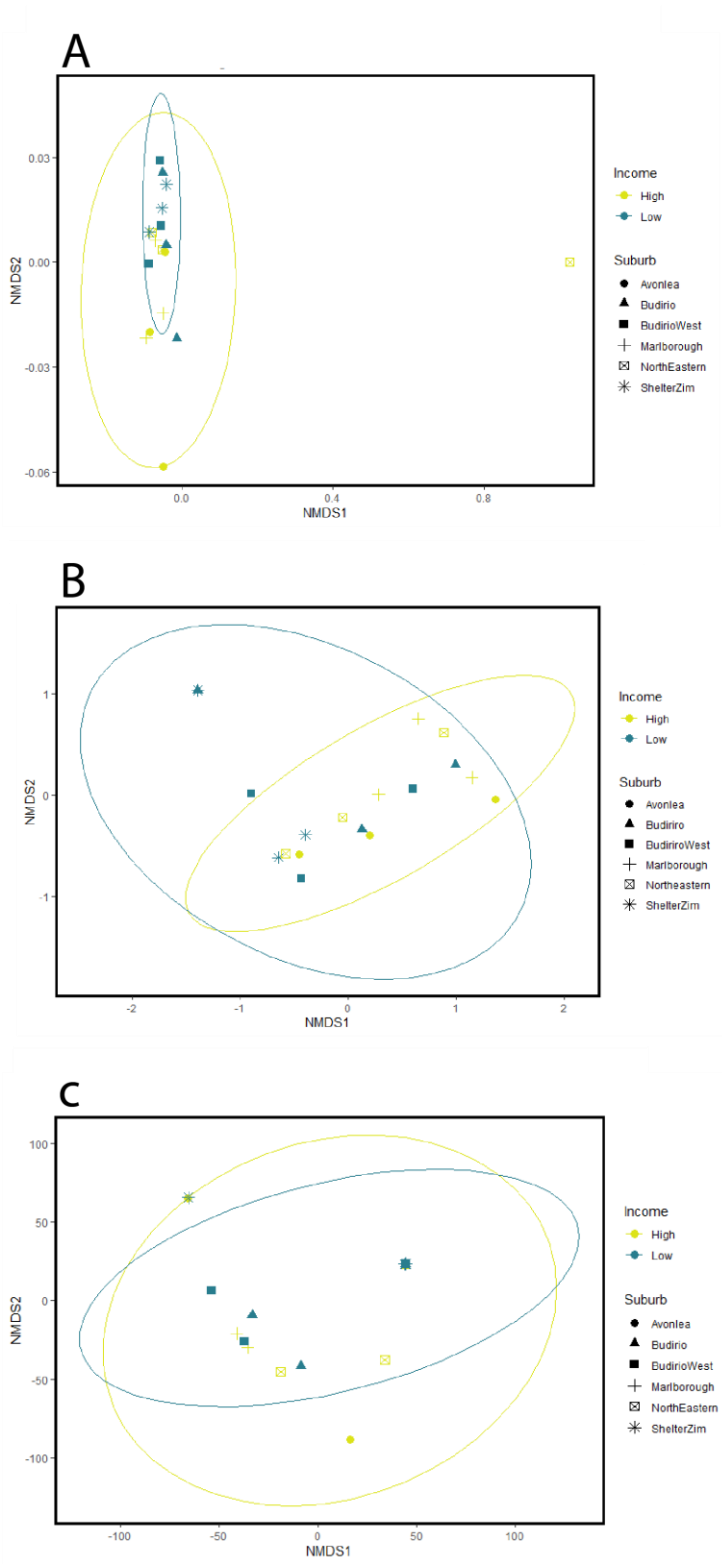
Supplementary Figure 7. Sankey plot breaking down virulence factor genes into unique classes and proteins: Includes the total number and count for each class and protein detected across all samples



Supplementary Figure 8. Relative Abundance of the Top 10 Genera in Wastewater Samples by Suburb and Collection Date: Relative abundance of the top 10 bacterial genera in wastewater samples by suburb and collection date. Genera not included in the top 10 were grouped into "Other".



Supplementary Figure 9. Comparison of Shannon Diversity Scores between High-income and Low-income Suburbs: A) Metagenomes B) Antibiotic Resistance Genes C) Virulence Factor Genes
Individual samples were not merged by respective suburb prior to data processing and bioinformatics.



Supplementary Figure 10. Comparison of Bray-Curtis Dissimilarity Distances between High-income and Low-income Suburbs: A) Metagenomes B) Antibiotic Resistance Genes C) Virulence Factor Genes. Individual samples were not merged by respective suburb prior to data processing and bioinformatics. Note: NMDS units are not consistent between each plot.

References

1. Abia, A. L. K., Baloyi, T., Traore, A. N., & Potgieter, N. (2023). The African Wastewater Resistome: Identifying Knowledge Gaps to Inform Future Research Directions. *Antibiotics*, *12*(5), 805. https://doi.org/10.3390/antibiotics12050805
2. Adebisi, Y. A., Jimoh, N. D., Ogunkola, I. O., Uwizeyimana, T., Olayemi, A. H., Ukor, N. A., & Lucero-Prisno, D. E. (2021). The use of antibiotics in COVID-19 management: A rapid review of national treatment guidelines in 10 African countries. *Tropical Medicine and Health*, *49*, 51. https://doi.org/10.1186/s41182-021-00344-w
3. Alcock, B. P., Huynh, W., Chalil, R., Smith, K. W., Raphenya, A. R., Wlodarski, M. A., Edalatmand, A., Petkau, A., Syed, S. A., Tsang, K. K., Baker, S. J. C., Dave, M., McCarthy, M. C., Mukiri, K. M., Nasir, J. A., Golbon, B., Imtiaz, H., Jiang, X., Kaur, K., … McArthur, A. G. (2023). CARD 2023: Expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Research*, *51*(D1), D690–D699. https://doi.org/10.1093/nar/gkac920
4. *Antimicrobial resistance*. (n.d.). Retrieved January 29, 2024, from https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance
5. Bartlett, A., Padfield, D., Lear, L., Bendall, R., & Vos, M. (2022). A comprehensive list of bacterial pathogens infecting humans. *Microbiology*, *168*(12), 001269. https://doi.org/10.1099/mic.0.001269
6. Cahill, S. T., Tarhonskaya, H., Rydzik, A. M., Flashman, E., McDonough, M. A., Schofield, C. J., & Brem, J. (2016). Use of ferrous iron by metallo-β-lactamases. *Journal of Inorganic Biochemistry*, *163*, 185–193. https://doi.org/10.1016/j.jinorgbio.2016.07.013
7. CDC. (2018, March 23). Notes from the Field: Typhoid Fever Outbreak – Hrare, Zimbabwe, October 2016-March 2017. *Morbidity and Mortality Weekly Report, 67*(11), 342-343. https://www.cdc.gov/mmwr/volumes/67/wr/mm6711a7.htm
8. CDC. (2022, October 5). *What Exactly is Antibiotic Resistance?* Centers for Disease Control and Prevention. https://www.cdc.gov/drugresistance/about.html
9. Chimbwanda, Tatenda. (2022, May 12). *High level of resistant bacteria circulating in the environment, food chain and community: A call to strengthen One Health Surveillance on antimicrobial resistance in Zimbabwe*. WHO | Regional Office for Africa. https://www.afro.who.int/countries/zimbabwe/news/high-level-resistant-bacteria-circulating-environment-food-chain-and-community-call-strengthen-one
10. Chirgwin, H., Cairncross, S., Zehra, D., & Sharma Waddington, H. (2021). Interventions promoting uptake of water, sanitation and hygiene (WASH) technologies in low- and middle-income countries: An evidence and gap map of effectiveness studies. *Campbell Systematic Reviews*, *17*(4), e1194. https://doi.org/10.1002/cl2.1194
11. Chirisa, I., Bandauko, E., Matamanda, A., & Mandisvika, G. (2017). Decentralized domestic wastewater systems in developing countries: The case study of Harare (Zimbabwe). *Applied Water Science*, *7*(3), 1069–1078. https://doi.org/10.1007/s13201-016-0377-4
12. Chitungo, I., Dzinamarira, T., Nyazika, T. K., Herrera, H., Musuka, G., & Murewanhema, G. (2022). Inappropriate Antibiotic Use in Zimbabwe in the COVID-19 Era: A Perfect Recipe for Antimicrobial Resistance. *Antibiotics*, *11*(2), 244. https://doi.org/10.3390/antibiotics11020244
13. Conner, J. G., Teschler, J. K., Jones, C. J., & Yildiz, F. H. (2016). Staying Alive: Vibrio cholerae's Cycle of Environmental Survival, Transmission, and Dissemination.

Microbiology Spectrum, *4*(2), 10.1128/microbiolspec.vmbf-0015–2015. https://doi.org/10.1128/microbiolspec.vmbf-0015-2015
14. Coulthurst, S. (2019). The Type VI secretion system: A versatile bacterial weapon. *Microbiology*, *165*(5), 503–515. https://doi.org/10.1099/mic.0.000789
15. Deng, W., Marshall, N. C., Rowland, J. L., McCoy, J. M., Worrall, L. J., Santos, A. S., Strynadka, N. C. J., & Finlay, B. B. (2017). Assembly, structure, function and regulation of type III secretion systems. *Nature Reviews Microbiology*, *15*(6), 323–337. https://doi.org/10.1038/nrmicro.2017.20
16. Fagnant, C. S., Beck, N. K., Yang, M.-F., Barnes, K. S., Boyle, D. S., & Meschke, J. S. (2014). Development of a novel bag-mediated filtration system for environmental recovery of poliovirus. Journal of Water and Health, 12(4), 747–754. https://doi.org/10.2166/wh.2014.032
17. Fagnant, C. S., Sánchez-Gonzalez, L. M., Zhou, N. A., Falman, J. C., Eisenstein, M., Guelig, D., Ockerman, B., Guan, Y., Kossik, A. L., Linden, Y. S., Beck, N. K., Wilmouth, R., Komen, E., Mwangi, B., Nyangao, J., Shirai, J. H., Novosselov, I., Borus, P., Boyle, D. S., & Meschke, J. S. (2018). Improvement of the Bag-Mediated Filtration System for Sampling Wastewater and Wastewater-Impacted Waters. Food and Environmental Virology, 10(1), 72–82. https://doi.org/10.1007/s12560-017-9311-7
18. Fagnant, C. S., Toles, M., Zhou, N. A., Powell, J., Adolphsen, J., Guan, Y., Ockerman, B., Shirai, J. H., Boyle, D. S., Novosselov, I., & Meschke, J. S. (2017). Development of an elution device for ViroCap virus filters. Environmental Monitoring and Assessment, 189(11). https://doi.org/10.1007/s10661-017-6258-y
19. Ferdous, F., Das, S. K., Ahmed, S., Farzana, F. D., Latham, J. R., Chisti, M. J., Ud-Din, A. I. M. S., Azmi, I. J., Talukder, K. A., & Faruque, A. S. G. (2013). Severity of Diarrhea and Malnutrition among Under Five-Year-Old Children in Rural Bangladesh. *The American Journal of Tropical Medicine and Hygiene*, *89*(2), 223–228. https://doi.org/10.4269/ajtmh.12-0743
20. Green, E. R., & Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiology Spectrum*, *4*(1), 10.1128/microbiolspec.vmbf-0012–2015. https://doi.org/10.1128/microbiolspec.vmbf-0012-2015
21. Gufe, C., Ndlovu, M. N., Sibanda, Z., Makuvara, Z., & Marumure, J. (2021). Prevalence and antimicrobial profile of potentially pathogenic bacteria isolated from abattoir effluents in Bulawayo, Zimbabwe. *Scientific African*, *14*, e01059. https://doi.org/10.1016/j.sciaf.2021.e01059
22. Gwatidzo, S. D., Murambinda, P. K., & Makoni, Z. (2017). Medicines Counterfeiting in Africa: A View from Zimbabwe. *Medicine Access @ Point of Care*, *1*, maapoc.0000017. https://doi.org/10.5301/maapoc.0000017
23. Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*(1), 119. https://doi.org/10.1186/1471-2105-11-119
24. IFRC. (2023). *DREF Operational Update.* IFRC.
25. Jankowski, P., Gan, J., Le, T., McKennitt, M., Garcia, A., Yanaç, K., Yuan, Q., & Uyaguari-Diaz, M. (2022). Metagenomic community composition and resistome analysis in a full-scale cold climate wastewater treatment plant. *Environmental Microbiome*, *17*, 3. https://doi.org/10.1186/s40793-022-00398-1

26. Jespersen, M. L., Munk, P., Johansen, J., Kaas, R. S., Webel, H., Vigre, H., Nielsen, H. B., Rasmussen, S., & Aarestrup, F. M. (2023). Global within-species phylogenetics of sewage microbes suggest that local adaptation shapes geographical bacterial clustering. *Communications Biology*, *6*(1), 1–9. https://doi.org/10.1038/s42003-023-05083-8
27. Kaper, J. B., Nataro, J. P., & Mobley, H. L. T. (2004). Pathogenic Escherichia coli. *Nature Reviews Microbiology*, *2*(2), 123–140. https://doi.org/10.1038/nrmicro818
28. Khan, F. Z., Palmer, K., & Dillon, N. (2024). Siderophores mediate antibiotic resistance. *Nature Microbiology*, *9*(3), 587–588. https://doi.org/10.1038/s41564-024-01624-x
29. Khazaal, M. T., El-hendawy, H. H., Mabrouk, M. I., Faraag, A. H. I., & Bakkar, M. R. (2022). Antibiotic resistance and siderophores production by clinical Escherichia coli strains. *BioTechnologia*, *103*(2), 169–184. https://doi.org/10.5114/bta.2022.116211
30. Kitakawa, K., Kitamura, K., & Yoshida, H. (n.d.). Monitoring Enteroviruses and SARS-CoV-2 in Wastewater Using the Polio Environmental Surveillance System in Japan. *Applied and Environmental Microbiology*, *89*(4), e01853-22. https://doi.org/10.1128/aem.01853-22
31. Korotkov, K. V., & Sandkvist, M. (2019). Architecture, function, and substrates of the type II secretion system. *EcoSal Plus*, *8*(2), 10.1128/ecosalplus.ESP-0034–2018. https://doi.org/10.1128/ecosalplus.ESP-0034-2018
32. Krawczyk, P. S., Lipinski, L., & Dziembowski, A. (2018). PlasFlow: Predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Research*, *46*(6), e35. https://doi.org/10.1093/nar/gkx1321
33. Kuhn, K. G., Shukla, R., Mannell, M., Graves, G. M., Miller, A. C., Vogel, J., Malloy, K., Deshpande, G., Florea, G., Shelton, K., Jeffries, E., De León, K. B., & Stevenson, B. (2023). Using Wastewater Surveillance to Monitor Gastrointestinal Pathogen Infections in the State of Oklahoma. *Microorganisms*, *11*(9), Article 9. https://doi.org/10.3390/microorganisms11092193
34. Lee, J., Ju, F., Beck, K., & Bürgmann, H. (2023). Differential effects of wastewater treatment plant effluents on the antibiotic resistomes of diverse river habitats. *The ISME Journal*, *17*(11), 1993–2002. https://doi.org/10.1038/s41396-023-01506-w
35. Lu J, Breitwieser FP, Thielen P, Salzberg SL. (2017). Bracken: estimating species abundance in metagenomics data. *PeerJ Computer Science* 3:e104, doi:10.7717/peerj-cs.104
36. Ma, Y., Wu, N., Zhang, T., Li, Y., Cao, L., Zhang, P., Zhang, Z., Zhu, T., & Zhang, C. (n.d.). The microbiome, resistome, and their co-evolution in sewage at a hospital for infectious diseases in Shanghai, China. *Microbiology Spectrum*, *12*(2), e03900-23. https://doi.org/10.1128/spectrum.03900-23
37. Mashe, T., Chaibva, B. V., Nair, P., Sani, K. A., Jallow, M., Tarupiwa, A., Goredema, A., Munyanyi, M., Chimusoro, A., Mpala, N., Masunda, K. P. E., Duri, C., Chonzi, P., & Phiri, I. (2023). Descriptive epidemiology of the cholera outbreak in Zimbabwe 2018–2019: Role of multi-sectorial approach in cholera epidemic control. *BMJ Open*, *13*(1), e059134. https://doi.org/10.1136/bmjopen-2021-059134
38. Mauchaza, K., Madzimbamuto, F. D., & Waner, S. (2016). Methicillin-resistant Staphylococcus aureus in Zimbabwe. *Ghana Medical Journal*, *50*(2), 68–71.
39. Mhondoro, M., Ndlovu, N., Bangure, D., Juru, T., Gombe, N. T., Shambira, G., Nsubuga, P., & Tshimanga, M. (2019). Trends in antimicrobial resistance of bacterial pathogens in

Harare, Zimbabwe, 2012–2017: A secondary dataset analysis. *BMC Infectious Diseases*, *19*(1), 746. https://doi.org/10.1186/s12879-019-4295-6

40. Moyo, N. A. G. (2013). An analysis of the chemical and microbiological quality of ground water from boreholes and shallow wells in Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C*, *66*, 27–32. https://doi.org/10.1016/j.pce.2013.06.003
41. Munk, P., Brinch, C., Møller, F. D., Petersen, T. N., Hendriksen, R. S., Seyfarth, A. M., Kjeldgaard, J. S., Svendsen, C. A., van Bunnik, B., Berglund, F., Larsson, D. G. J., Koopmans, M., Woolhouse, M., & Aarestrup, F. M. (2022). Genomic analysis of sewage from 101 countries reveals global landscape of antimicrobial resistance. *Nature Communications*, *13*(1), 7251. https://doi.org/10.1038/s41467-022-34312-7
42. Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). metaSPAdes: A new versatile metagenomic assembler. *Genome Research*, *27*(5), 824–834. https://doi.org/10.1101/gr.213959.116
43. Olaru, I. D., Ferrand, R. A., Yeung, S., Chingono, R., Chonzi, P., Masunda, K. P. E., Dixon, J., & Kranzer, K. (2022b). Knowledge, attitudes and practices relating to antibiotic use and resistance among prescribers from public primary healthcare facilities in Harare, Zimbabwe. *Wellcome Open Research*, *6*, 72. https://doi.org/10.12688/wellcomeopenres.16657.2
44. Parween, F., Yadav, J., & Qadri, A. (2019). The Virulence Polysaccharide of Salmonella Typhi Suppresses Activation of Rho Family GTPases to Limit Inflammatory Responses From Epithelial Cells. *Frontiers in Cellular and Infection Microbiology*, *9*. https://doi.org/10.3389/fcimb.2019.00141
45. Pavlinac, P. B., Tickell, K. D., & Walson, J. L. (2015). Management of diarrhea in HIV-affected infants and children. *Expert Review of Anti-Infective Therapy*, *13*(1), 5–8. https://doi.org/10.1586/14787210.2015.981157
46. Rajeev, M., Jung, I., Lim, Y., Kim, S., Kang, I., & Cho, J.-C. (2023). Metagenome sequencing and recovery of 444 metagenome-assembled genomes from the biofloc aquaculture system. *Scientific Data*, *10*, 707. https://doi.org/10.1038/s41597-023-02622-0
47. Ramees, T. P., Dhama, K., Karthik, K., Rathore, R. S., Kumar, A., Saminathan, M., Tiwari, R., Malik, Y. S., & Singh, R. K. (2017). Arcobacter: An emerging food-borne zoonotic pathogen, its public health concerns and advances in diagnosis and control – a comprehensive review. *Veterinary Quarterly*, *37*(1), 136–161. https://doi.org/10.1080/01652176.2017.1323355
48. Rodriguez-R LM & Konstantinidis KT (2016). The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. https://peerj.com/preprints/1900/.
49. Rout, A. K., Tripathy, P. S., Dixit, S., Behera, D. U., Behera, B., Das, B. K., & Behera, B. K. (2023). Unveiling the Microbiome Landscape: A Metagenomic Study of Bacterial Diversity, Antibiotic Resistance, and Virulence Factors in the Sediments of the River Ganga, India. *Antibiotics*, *12*(12), 1735. https://doi.org/10.3390/antibiotics12121735
50. Ruhanya, V., Zhou, N. A., Berejena, C., Nyandoro, G., Chibukira, P., Mukaratirwa, A., Muserere, S. T., Masunda, K., & Meschke, J. S. (2022). Recovery of enteroviruses and poliovirus in Harare sewage using the bag-mediated filtration system at the introduction of the inactivated polio vaccine in Zimbabwe. *Water SA*, *48*(3), 248–254. https://doi.org/10.17159/wsa/2022.v48.i3.3892

51. Sartorius, B., Gray, A. P., Weaver, N. D., Aguilar, G. R., Swetschinski, L. R., Ikuta, K. S., Mestrovic, T., Chung, E., Wool, E. E., Han, C., Hayoon, A. G., Araki, D. T., Abd-Elsalam, S., Aboagye, R. G., Adamu, L. H., Adepoju, A. V., Ahmed, A., Akalu, G. T., Akande-Sholabi, W., … Naghavi, M. (2024). The burden of bacterial antimicrobial resistance in the WHO African region in 2019: A cross-country systematic analysis. *The Lancet Global Health*, *12*(2), e201–e216. https://doi.org/10.1016/S2214-109X(23)00539-9
52. Schneeberger, P. H. H., Fuhrimann, S., Becker, S. L., Pothier, J. F., Duffy, B., Beuret, C., Frey, J. E., & Utzinger, J. (2019). Qualitative microbiome profiling along a wastewater system in Kampala, Uganda. *Scientific Reports*, *9*, 17334. https://doi.org/10.1038/s41598-019-53569-5
53. Sharma, A. K., Dhasmana, N., Dubey, N., Kumar, N., Gangwal, A., Gupta, M., & Singh, Y. (2017). Bacterial Virulence Factors: Secreted for Survival. *Indian Journal of Microbiology*, *57*(1), 1–10. https://doi.org/10.1007/s12088-016-0625-1
54. Sharma, J., Sharma, D., Singh, A., & Sunita, K. (2022). Colistin Resistance and Management of Drug Resistant Infections. *The Canadian Journal of Infectious Diseases & Medical Microbiology = Journal Canadien Des Maladies Infectieuses et de La Microbiologie Médicale*, *2022*, 4315030. https://doi.org/10.1155/2022/4315030
55. Shen, J., McFarland, A. G., Young, V. B., Hayden, M. K., & Hartmann, E. M. (2021). Toward Accurate and Robust Environmental Surveillance Using Metagenomics. *Frontiers in Genetics*, *12*. https://www.frontiersin.org/articles/10.3389/fgene.2021.600111
56. Solís-Sánchez, Pablo & Fernández-Martínez, Marta & Rodrigo-Calabia, Emilio & Alegría-Puig, Carlos. (2023). Chronic Diarrhea Due to Aeromonas hydrophila in an Immunosuppressed Patient with a Pancreas–Kidney Transplant. Pathogens. 12. 1151. 10.3390/pathogens12091151.
57. Street, R., Malema, S., Mahlangeni, N., & Mathee, A. (2020). Wastewater surveillance for Covid-19: An African perspective. *The Science of the Total Environment*, *743*, 140719. https://doi.org/10.1016/j.scitotenv.2020.140719
58. Takawira, F. T., Pitout, J. D. D., Thilliez, G., Mashe, T., Gutierrez, A. V., Kingsley, R. A., Peirano, G., Matheu, J., Midzi, S. M., Mwamakamba, L. W., Gally, D. L., Tarupiwa, A., Mukavhi, L., Ehlers, M. M., Mtapuri-Zinyowera, S., & Kock, M. M. (2022). Faecal carriage of ESBL producing and colistin resistant Escherichia coli in avian species over a 2-year period (2017-2019) in Zimbabwe. *Frontiers in Cellular and Infection Microbiology*, *12*, 1035145. https://doi.org/10.3389/fcimb.2022.1035145
59. Tiwari, A., Hokajärvi, A.-M., Domingo, J. S., Elk, M., Jayaprakash, B., Ryu, H., Siponen, S., Vepsäläinen, A., Kauppinen, A., Puurunen, O., Artimo, A., Perkola, N., Huttula, T., Miettinen, I. T., & Pitkänen, T. (2021). Bacterial diversity and predicted enzymatic function in a multipurpose surface water system – from wastewater effluent discharges to drinking water production. *Environmental Microbiome*, *16*, 11. https://doi.org/10.1186/s40793-021-00379-w
60. Tsiko, R., & Togarepi, S. (2012). A Situational Analysis of Waste Management in Harare, Zimbabwe. *American Journal of Science*, *8(4)*, 692–706.
61. UNICEF. (2023, July 19). *Ending HIV/AIDS with Children, Adolescents and Young Women.* UNICEF Zimbabwe. https://www.unicef.org/zimbabwe/reports/ending-hivaids-children-adolescents-and-young-women

62. Uzzell, C. B., Abraham, D., Rigby, J., Troman, C. M., Nair, S., Elviss, N., Kathiresan, L., Srinivasan, R., Balaji, V., Zhou, N. A., Meschke, J. S., John, J., Kang, G., Feasey, N., Mohan, V. R., & Grassly, N. C. (2024). Environmental Surveillance for Salmonella Typhi and its Association With Typhoid Fever Incidence in India and Malawi. *The Journal of Infectious Diseases*, *229*(4), 979–987. https://doi.org/10.1093/infdis/jiad427
63. WHO. (2018, October 3). *Zimbabwe to vaccinate 1.4 million people against cholera in Harare.* WHO. https://www.afro.who.int/news/zimbabwe-vaccinate-14-million-people-against-cholera-harare
64. WHO. (2018, October 5). *Cholera – Zimbabwe*. WHO. https://www.who.int/emergencies/disease-outbreak-news/item/05-october-2018-cholera-zimbabwe-en#:~:text=As%20of%203%20October%202018,capital%20Harare%20(Figure%201).
65. WHO. (2023, June 28). *Improving access to water, sanitation and hygiene can save 1.4 million lives per year, says new WHO report.* WHO. https://www.who.int/news/item/28-06-2023-improving-access-to-water--sanitation-and-hygiene-can-save-1.4-million-lives-per-year--says-new-who-report
66. WHO. (2024, March 7). *Diarrhoeal disease*. WHO. https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease
67. Zhang, B., Xia, Y., Wen, X., Wang, X., Yang, Y., Zhou, J., & Zhang, Y. (2016). The Composition and Spatial Patterns of Bacterial Virulence Factors and Antibiotic Resistance Genes in 19 Wastewater Treatment Plants. *PLoS ONE*, *11*(12), e0167422. https://doi.org/10.1371/journal.pone.0167422
68. Zhou, N. A., Fagnant-Sperati, C. S., Shirai, J. H., Sharif, S., Zaidi, S. Z., Rehman, L., Hussain, J., Agha, R., Shaukat, S., Alam, M., Khurshid, A., Mujtaba, G., Salman, M., Safdar, R. M., Mahamud, A., Ahmed, J., Khan, S., Kossik, A. L., Beck, N. K., … Meschke, J. S. (2018). Evaluation of the bag-mediated filtration system as a novel tool for poliovirus environmental surveillance: Results from a comparative field study in Pakistan. *PLOS ONE*, *13*(7), e0200551. https://doi.org/10.1371/journal.pone.0200551

Additional References

1. Aaltonen, K., Kant, R., Eklund, M., Raunio-Saarnisto, M., Paulin, L., Vapalahti, O., Grönthal, T., Rantala, M., & Sironen, T. (2020). Streptococcus halichoeri: Comparative Genomics of an Emerging Pathogen. *International Journal of Genomics*, *2020*, 8708305. https://doi.org/10.1155/2020/8708305
2. *Alignment*. (n.d.). Retrieved November 10, 2023, from https://support.illumina.com/help/BS_App_TS_Amplicon_v3_OLH_15055858/Content/Source/Informatics/Alignment_SmithWaterman.htm
3. Antimicrobial Resistance. *Microbiology Spectrum*, *11*(2), e04102-22. https://doi.org/10.1128/spectrum.04102-22
4. Asghar, H., Diop, O. M., Weldegebriel, G., Malik, F., Shetty, S., Bassioni, L. E., Akande, A. O., Maamoun, E. A., Zaidi, S., Adeniji, A. J., Burns, C. C., Deshpande, J., Oberste, M. S., & Lowther, S. A. (2014). Environmental Surveillance for Polioviruses in the Global Polio Eradication Initiative. *The Journal of Infectious Diseases*, *210*(Suppl 1), S294–S303. https://doi.org/10.1093/infdis/jiu384
5. Azli, B., Razak, M. N., Omar, A. R., Mohd Zain, N. A., Abdul Razak, F., & Nurulfiza, I. (2022). Metagenomics Insights Into the Microbial Diversity and Microbiome Network Analysis on the Heterogeneity of Influent to Effluent Water. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.779196
6. Brumfield, K. D., Leddy, M., Usmani, M., Cotruvo, J. A., Tien, C.-T., Dorsey, S., Graubics, K., Fanelli, B., Zhou, I., Registe, N., Dadlani, M., Wimalarante, M., Jinasena, D., Abayagunawardena, R., Withanachchi, C., Huq, A., Jutla, A., & Colwell, R. R. (2022.). Microbiome Analysis for Wastewater Surveillance during COVID-19. *mBio*, *13*(4), e00591-22. https://doi.org/10.1128/mbio.00591-22
7. Chari, S., Mbonane, T. P., & Van Wyk, R. H. (2023). Social and Environmental Determinants of Diarrheal Diseases among Children under Five Years in Epworth Township, Harare. *Children*, *10*(7), 1173. https://doi.org/10.3390/children10071173
8. Douidah, L., de Zutter, L., Baré, J., De Vos, P., Vandamme, P., Vandenberg, O., Van den Abeele, A.-M., & Houf, K. (2012). Occurrence of Putative Virulence Genes in Arcobacter Species Isolated from Humans and Animals. *Journal of Clinical Microbiology*, *50*(3), 735–741. https://doi.org/10.1128/JCM.05872-11
9. Edgar, R. C., & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, *31*(21), 3476–3482. https://doi.org/10.1093/bioinformatics/btv401
10. Einstein, G. P., Tulp, O. L., Oyesile, T. C., & Obidi, O. F. (2017). Meta-analysis of staphylococcal diarrhea in some developing African countries. *Beni-Suef University Journal of Basic and Applied Sciences*, *6*(2), 145–148. https://doi.org/10.1016/j.bjbas.2017.02.004
11. Gumbo, A., Bangure, D., Gombe, N. T., Mungati, M., Tshimanga, M., Hwalima, Z., & Dube, I. (2015). Staphylococcus aureus food poisoning among Bulawayo City Council employees, Zimbabwe, 2014. *BMC Research Notes*, *8*, 485. https://doi.org/10.1186/s13104-015-1490-4
12. Hong, J.-K., Oh, H., Lee, T. K., Kim, S., Oh, D., Ahn, J., & Park, S. (2023). Tracking the Evolution of Microbial Communities on Microplastics through a Wastewater Treatment Process: Insight into the "Plastisphere." *Water*, *15*(21), Article 21. https://doi.org/10.3390/w15213746

13. Chimbwanda, Tatenda. (2022, April 11). *Increasing momentum to address Antimicrobial Resistance in Zimbabwe: Commitment of the government and the UN Country Team*. WHO | Regional Office for Africa. https://www.afro.who.int/countries/zimbabwe/news/increasing-momentum-address-antimicrobial-resistance-zimbabwe-commitment-government-and-un-country
14. Kers, Jannigje Gerdien, and Edoardo Saccenti. "The Power of Microbiome Studies: Some Considerations on Which Alpha and Beta Metrics to Use and How to Report Results." *Frontiers in Microbiology* 12 (2022). https://www.frontiersin.org/articles/10.3389/fmicb.2021.796025.
15. Kietsiri, P., Muangnapoh, C., Lurchachaiwong, W., Lertsethtakarn, P., Bodhidatta, L., Suthienkul, O., Waters, N. C., Demons, S. T., & Vesely, B. A. (2021). Characterization of Arcobacter spp. Isolated from human diarrheal, non-diarrheal and food samples in Thailand. *PLOS ONE*, *16*(2), e0246598. https://doi.org/10.1371/journal.pone.0246598
16. LaMartina, E. L., Mohaimani, A. A., & Newton, R. J. (2021). Urban wastewater bacterial communities assemble into seasonal steady states. *Microbiome*, *9*(1), 116. https://doi.org/10.1186/s40168-021-01038-5
17. Liao, C., Huang, X., Wang, Q., Yao, D., & Lu, W. (2022). Virulence Factors of Pseudomonas Aeruginosa and Antivirulence Strategies to Combat Its Drug Resistance. *Frontiers in Cellular and Infection Microbiology*, *12*. https://doi.org/10.3389/fcimb.2022.926758
18. Lisiecki, Pawel. (2014). Antibiotic resistance and siderophore production in enterococci. Medycyna doświadczalna i mikrobiologia. 66. 1-10.
19. Liu, B., Zheng, D., Zhou, S., Chen, L., & Yang, J. (2021). VFDB 2022: A general classification scheme for bacterial virulence factors. *Nucleic Acids Research*, *50*(D1), D912–D917. https://doi.org/10.1093/nar/gkab1107
20. Madamombe, K., Karakadzai, M., Masoja, G., Dhliwayo, T., Juru, T., Chadambuka, A., Govha, E., Gombe, N., & Tshimanga, M. (2022). Evaluation of the weekly disease surveillance system for epidemic-prone diseases in Makonde District, Zimbabwe 2020: A descriptive cross-sectional study. *The Pan African Medical Journal*, *43*, 132. https://doi.org/10.11604/pamj.2022.43.132.35001
21. Marutescu, L. G., Popa, M., Gheorghe-Barbu, I., Barbu, I. C., Rodríguez-Molina, D., Berglund, F., Blaak, H., Flach, C.-F., Kemper, M. A., Spießberger, B., Wengenroth, L., Larsson, D. G. J., Nowak, D., Radon, K., de Roda Husman, A. M., Wieser, A., Schmitt, H., Pircalabioru Gradisteanu, G., Vrancianu, C. O., & Chifiriuc, M. C. (2023). Wastewater treatment plants, an "escape gate" for ESCAPE pathogens. *Frontiers in Microbiology*, *14*. https://doi.org/10.3389/fmicb.2023.1193907
22. Muserere, S. T., Hoko, Z., & Nhapi, I. (2014). Characterisation of raw sewage and performance assessment of primary settling tanks at Firle Sewage Treatment Works, Harare, Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C*, *67–69*, 226–235. https://doi.org/10.1016/j.pce.2013.10.004
23. Musuka, G., Dzinamarira, T., Murewanhema, G., Cuadros, D., Chingombe, I., Herrera, H., Takavarasha, F., & Mapingure, M. (2021). Associations of diarrhea episodes and seeking medical treatment among children under five years: Insights from the Zimbabwe Demographic Health Survey (2015–2016). *Food Science & Nutrition*, *9*(11), 6335–6342. https://doi.org/10.1002/fsn3.2596

24. Nhapi, Innocent. (2004). Options for wastewater management in Harare, Zimbabwe. https://edepot.wur.nl/193572
25. Oneko, M., Nyathi, M. N., & Doehring, E. (2001). Post-dysenteric hemolytic uremic syndrome in Bulawayo, Zimbabwe. *Pediatric Nephrology*, *16*(12), 1142–1145. https://doi.org/10.1007/s004670100049
26. O'Ryan, M. L., Nataro, J. P., & Cleary, T. G. (2011). Microorganisms Responsible for Neonatal Diarrhea. *Infectious Diseases of the Fetus and Newborn*, 359–418. https://doi.org/10.1016/B978-1-4160-6400-8.00011-0
27. Plaza, G., Jałowiecki, Ł., Głowacka, D., Hubeny, J., Harnisz, M., & Korzeniewska, E. (2021). Insights into the microbial diversity and structure in a full-scale municipal wastewater treatment plant with particular regard to Archaea. *PLoS ONE*, *16*(4), e0250514. https://doi.org/10.1371/journal.pone.0250514
28. Portik, D. M., Brown, C. T., & Pierce-Ward, N. T. (2022). Evaluation of taxonomic classification and profiling methods for long-read shotgun metagenomic sequencing datasets. *BMC Bioinformatics*, *23*(1), 541. https://doi.org/10.1186/s12859-022-05103-0
29. Roguet, A., Newton, R. J., Eren, A. M., & McLellan, S. L. (2022). Guts of the Urban Ecosystem: Microbial Ecology of Sewer Infrastructure. *mSystems*, *7*(4), e00118-22. https://doi.org/10.1128/msystems.00118-22
30. Salazar-Sánchez, A., Baztarrika, I., Alonso, R., Fernández-Astorga, A., Martínez-Ballesteros, I., & Martinez-Malaxetxebarria, I. (2022). Arcobacter butzleri Biofilms: Insights into the Genes Beneath Their Formation. *Microorganisms*, *10*(7), 1280. https://doi.org/10.3390/microorganisms10071280
31. Scaturro, M., Del Chierico, F., Motro, Y., Chaldoupi, A., Flountzi, A., Moran-Gilad, J., Girolamo, A., Koutsiomani, T., Krogulska, B., Lindsay, D., Matuszewska, R., Papageorgiou, G., Pancer, K., Panoussis, N., Rota, M. C., Uldum, S. A., Velonakis, E., Chaput, D. L., & Ricci, M. L. (2023). Premise plumbing bacterial communities in four European cities and their association with Legionella. *Frontiers in Microbiomes*, *2*. https://doi.org/10.3389/frmbi.2023.1170824
32. Shah, C., Baral, R., Bartaula, B., & Shrestha, L. B. (2019). Virulence factors of uropathogenic Escherichia coli (UPEC) and correlation with antimicrobial resistance. *BMC Microbiology*, *19*, 204. https://doi.org/10.1186/s12866-019-1587-3
33. Standard, T. (n.d.). *Budiriro, a constituency waiting on their MP*. NewsDay. Retrieved April 29, 2024, from https://www.newsday.co.zw/news/article/74150/budiriro-a-constituency-waiting-on-their-mp
34. Talat, A., Blake, K. S., Dantas, G., & Khan, A. U. (2023). Metagenomic Insight into Microbiome and Antibiotic Resistance Genes of High Clinical Concern in Urban and Rural Hospital Wastewater of Northern India Origin: A Major Reservoir of
35. Tan, B., Ng, C., Nshimyimana, J., Loh, L.-L., Gin, K., & Thompson, J. (2015). Next-generation sequencing (NGS) for assessment of microbial water quality: Current progress, challenges, and future opportunities. *Frontiers in Microbiology*, *6*. https://www.frontiersin.org/articles/10.3389/fmicb.2015.01027
36. Vega, L., Jaimes, J., Morales, D., Martínez, D., Cruz-Saavedra, L., Muñoz, M., & Ramírez, J. D. (2021). Microbial Communities' Characterization in Urban Recreational Surface Waters Using Next Generation Sequencing. *Microbial Ecology*, *81*(4), 847–863. https://doi.org/10.1007/s00248-020-01649-9

37. Wang, J., Chen, Y., Cai, P., Gao, Q., Zhong, H., Sun, W., & Chen, Q. (2022). Impacts of municipal wastewater treatment plant discharge on microbial community structure and function of the receiving river in Northwest Tibetan Plateau. *Journal of Hazardous Materials*, *423*, 127170. https://doi.org/10.1016/j.jhazmat.2021.127170
38. Wang, Z., Zhang, X.-X., Huang, K., Miao, Y., Shi, P., Liu, B., Long, C., & Li, A. (2013). Metagenomic Profiling of Antibiotic Resistance Genes and Mobile Genetic Elements in a Tannery Wastewater Treatment Plant. *PLoS ONE*, *8*(10), e76079. https://doi.org/10.1371/journal.pone.0076079
39. Weldegebriel, G. (2015). Environmental Surveillance for Poliovirus in Polio High Risk States of Nigeria, 2011- 2012. *Science Journal of Public Health*, *3*(5), 655. https://doi.org/10.11648/j.sjph.20150305.20
40. Chimbwanda, Tatenda. (2020, December 14). *Zimbabwe's Antimicrobial Resistance threat to grow if not tackled in a holistic manner*. WHO | Regional Office for Africa. https://www.afro.who.int/countries/zimbabwe/news/zimbabwes-antimicrobial-resistance-threat-grow-if-not-tackled-holistic-manner