5-2018

Antibiotic resistant bacteria in water environments in Louisville, Kentucky.

Amy Priest
University of Louisville

Antibiotic Resistant Bacteria in Water Environments in Louisville, Kentucky: An

Analysis of Common Genera and Community Diversity

By

Amy Priest

Submitted in partial fulfillment of the requirements for Graduation *summa cum laude* and for

Graduation with Honors from the Department of Biology

University of Louisville

May, 2018

Table of Contents

Abstract

   Antibiotic resistant bacterial strains are an increasing problem, particularly in clinical health care settings. As a result, bacterial infections are becoming increasingly challenging to treat with more cases becoming life threatening. Aquatic environments facilitate microbial diversity and the transfer of genetic elements and thus may serve as a reservoir for antibiotic resistant microbes. Human misuse of antibiotics may further facilitate the spread of resistance in water environments. With little known about the bacteria communities in local water environments, this study aimed to learn more about these populations through the following aims: 1) identify the microbial community composition from water environments around Louisville, KY; and 2) examine of the communities were resistant to two clinically used antibiotics—vancomycin and colistin. In this study, water sites were sampled and sorted into 4 categories: agricultural waters, commercial drains, natural waters, and wastewaters. In total, 155 single colony isolates resistant to vancomycin and colistin were identified through 16S sequencing. Whole community metagenomics analysis characterized the bacterial composition of 87 communities from the initial sample collection. Community diversity and the relationship between diversity and income was analyzed. One of the most striking results was the presence of *Ochrobactrum* sp. in 78 of the 87 communities. Two of the most prevalent genera, *Ochrobactrum* and *Pseudochrobactrum*, were characterized by assessing relative antibiotic resistance profiles and were found to be tolerant to high doses of a spectrum of antibiotics. Finally, a representative *Ochrobactrum* sp. isolate was tested for its ability to confer antibiotic resistance to a susceptible recipient bacterium. This *Ochrobactrum* sp. isolate was unable to transfer colistin resistance to another bacterial species, *Pseudomonas aeruginosa*, despite repeated efforts. The results indicate that there is a large diversity of microbes resistant to vancomycin and colistin though the ability of these microbes to transfer this resistance remains to be seen.

Introduction

Antibiotic Resistance

Since the introduction of antibiotics in 1928 with the discovery of penicillin, treating bacterial infections became much easier. Since the onset of clinical antibiotics, human life-expectancy has increased over 20 years [1]. However, in the decade following the introduction of antibiotics, resistant bacterial strains were already being isolated. In the present day, despite the production of many novel antibiotics, the threat of antibiotic resistance is a growing concern to public health [2,3]. The Centers for Disease Control and Prevention (CDC) estimated that antibiotic resistant bacteria are responsible for nearly 2 million infectious diseases and a subsequent 23,000 deaths each year in the United States [73].

Keeping some antibiotics on reserve as last resort antibiotics is imperative in treating multi-drug resistant pathogens. For example, vancomycin has long been considered a last resort antibiotic that has been particularly effective in treating infections from methicillin-resistant *Staphylococcus aureus* (MRSA). This pathogen is notorious for infecting hospitalized patients and is the leading cause of death by a single infectious agent [4, 5]. Colistin has also been cited as a last resort antibiotic and is important in treating illnesses caused by Gram-negative organisms. However, even last resort antibiotics, like vancomycin and colistin, are becoming ineffective against this pathogen [5, 23]. Additionally, carbapenem antibiotics have also been cited as last resort antibiotics, but rising cases of carbapenem-resistant Enterobacteriaceae threaten the potency of antibiotics in this class [6].

Aquatic Environments and Bacterial Diversity

With water covering 70% of Earth's surface, aquatic areas are attractive habitats with many diverse niches for bacteria to occupy. In particular, freshwater sources can provide high quantities of nutrients, such as higher carbon, nitrogen, phosphorus, and sulfur inputs that help facilitate an environment conducive for microbial growth without the osmotic stresses that saltwater can impose upon cells [7-10]. Because of the wide variety of nutrients, abiotic factors (such as pH, O_2 concentrations, UV, temperatures), and types of predators, small freshwater systems would be expected to have a high level of bacterial biodiversity [7–9]. Previous studies of microbial diversity in freshwater sites have primarily focused on geographically constrained sampling sites. A few studies have suggested that bacteria are found in ranges of 10^4–10^8 CFU/mL in freshwater habitats [7–9]. However, it is unknown as to whether the trends in biodiversity observed in other geographical locations are universally applicable.

Aquatic environments are also a hub for the transfer of genetic material between microbes through horizontal gene transfer [10]. In aquatic niches where bacterial concentrations are high, bacteria can transfer mobile genetic elements such as transposons, plasmids, and transmissible genetic islands between bacterial species and strains through the processes of transformation, conjugation, or transduction. This transfer can involve the exchange of genes that confer the ability to resist antibiotics which is especially important because many environmental microbes produce secondary metabolites such as antibiotics to kill competing neighbors. This leads to an evolutionary pressure to acquire and maintain antibiotic resistance in natural environments. The transfer of genetic elements between microbes in freshwaters may facilitate the spread of traits that are undesirable to humans, such as antibiotic resistance, in microbial populations [11].

Human-Mediated Spread of Resistance in Natural Environments

Though aquatic environments may facilitate the spread of natural resistance amongst bacteria, human overuse and misuse of exogenous antibiotics are also prominent factors in both selection for acquired bacterial antibiotic resistance and its spread through environmental biomes [12–15]. In 2010, an estimated 506 antibiotic prescriptions were written for every 1000 people in the United States, equating to approximately 154 million prescriptions a year. Of those prescriptions, 30% percent were estimated to be "inappropriate" — meaning, these prescriptions were either not needed or not the proper antibiotic to treat the patient's diagnosis [14]. Kentucky, along with seven neighboring Midwestern and southern states, was estimated to have the highest rates of antibiotic prescriptions in the United States, with a range of 996 – 1,237 prescriptions per 1000 people [73].

The widespread human use of antibiotics leads to the human misuse of antibiotics. For example, individuals may start a round of antibiotics, experience a relief of symptoms, and cease taking the antibiotic prematurely. However, some bacteria may still reside in the host and were only exposed to low levels of that antibiotic. This low-level exposure can lead to the selection for resistant bacteria or the expression of genes involved in antibiotic resistance if they are present but unexpressed in the genome [15]. Additionally, exposure to antibiotics can induce a bacterial stress response, the SOS response, which can result in a recombination of genes contained in integrons that will code for resistance [16]. Humans may also contribute to the environmental spread of antibiotic resistant bacteria by excreting portions of unmodified antibiotics, resistant bacteria themselves, or by disposing unused antibiotics in toilets. These products may then enter sewer systems and contaminate waste waters.

Another major contributor to the rise in antibiotic resistant bacteria stems from excessive use of antibiotics in agriculture. Approximately 80% of antibiotics produced in the United States are administered to animals or used in fertilizer [12]. It has been estimated that 75–90% of the antibiotics administered to the animals are excreted almost completely unmodified, adding another selective pressure for antibiotic resistance in soil and groundwater [17]. Water sources, particularly those utilized for commercial fish production, are also breeding grounds for resistance as antibiotics are often included in fish food. Additionally, contaminated water sources can lead to the spread of antibiotic resistant bacteria as they are often not constrained to a single geographic area [13].

Vancomycin and Colistin

The spread of antibiotic resistant bacteria includes last resort antibiotics, such as vancomycin and colistin—two important antibiotics with differing modes of action. Vancomycin is a glycopeptide antibiotic that disrupts the cell wall by inhibiting proper peptidoglycan formation. The antibiotic targets D-alanine residues in peptides, preventing the cross linking of the peptide chains that are attached to *N*-acetylmuramic acid (NAM), one of the two main backbone carbohydrates in peptidoglycan [18]. Because this antibiotic inhibits peptidoglycan maturation, it is particularly effective against Gram-positive pathogens. Inherent vancomycin resistance is possible in some cases where the antibiotic is too large to penetrate the peptidoglycan layer, or when D-alanine is not naturally used in the peptidoglycan peptide chains. Acquired modes of plasmid–mediated resistance are also possible. In one such resistance mechanism, abnormal peptidoglycan synthesis occurs, resulting in one of the normal terminal D–alanine residues (recognized by vancomycin) being replaced with lactate, which prevents vancomycin from binding and obstructing cross-link formation [19]. In another acquired mechanism of vancomycin

resistance, the D-alanine–D-alanine peptide in peptidoglycan is modified to become D-alanine–D-serine, which also reduces the affinity of vancomycin to these peptide chains [20].

Colistin is a polymyxin antibiotic in the polypeptide class of antibiotics. Colistin targets lipopolysaccharide (LPS), the component of Gram-negative outer membrane. Colistin's positively charged peptide chain forms an electrostatic bond with the negatively charged LPS. This binding disrupts the membrane structure and stability, resulting in leakage of cellular contents and ultimately cell death [22]. Because colistin targets the LPS, it is used primarily to treat infections caused by Gram-negative bacteria. For years, colistin was not used due to possible nephrotoxicity in patients. However, the rise of Gram-negative multi-drug resistant bacteria has made colistin an attractive treatment option today [21]. Due to the low utilization of colistin in the past, very few instances of colistin resistance have been reported except for rare mutations that often decreased the fitness of the host bacterium and were not transmittable to other bacterial hosts. However, recently, plasmid-mediated colistin resistance via the *mcr-1* gene was reported and is of great concern due to its transferability and stability once the plasmid is incorporated into the host DNA [23].

Impoverished Areas and Antibiotic Resistance

Though antibiotic resistance is a problem everywhere, previous studies have sparked concern as to whether impoverished areas may be hotspots for the development of antibiotic resistance [24,25]. Individuals with a limited income may be more likely to misuse antibiotics in attempts to save money. For example, some individuals may prematurely end antibiotic cycles and save leftovers for other illnesses (for which those antibiotics may not be appropriate or sufficient) or share antibiotics with a friend or family member in need [24,25]. Other studies have characterized antibiotic resistant bacterial communities in lower-income developing countries [26,

27]. However, very little is known about antibiotic resistant bacterial communities in lower income areas in the United States [27].

Project Rationale

This study began as an upper division research-based course, BIOL 501 "Microbial Ecology of Antibiotic Resistance" which was comprised 21 undergraduate and post–bac students. The students chose to study freshwater environments across Louisville, KY. Freshwater environments were selected because of their abundance in nutrients and abiotic factors which were hypothesized to contain a breadth of microbial growth. Additionally, it was hypothesized that freshwater sites may facilitate genetic exchange and be reservoirs for conferring antibiotic resistance. To study local freshwater environments, sampling sites were selected according to the following habitats: 1) natural waters, including lakes, rivers, and streams; 2) agricultural waters, including standing water and ponds on commercial farms; 3) commercial drains, including drain samples obtained from local fast–food restaurants and gas stations; 4) wastewaters, including water from storm drains and ditches. Students formed four teams to sample each habitat by collecting water samples or by swabbing faucets and drains to grow both on control and on antibiotic plates. Antibiotic plates contained vancomycin and colistin, in addition to a nutrient medium. The students chose to study vancomycin and colistin because of their clinical relevance as last resort antibiotics and their differing modes of action against host cells.

The main objective of the first stage of the project was to isolate antibiotic resistant bacteria in environmental water sources around Louisville and to learn how those different sources differed in antibiotic resistant community composition. A second stage of the project was then initiated with 6 students. The aim of the second stage of the project was to analyze whole communities

from freshwaters by quantifying bacterial diversity through measuring levels of alpha and beta diversity calculations, and comparing corresponding selective and non-selective communities. Finally, the third stage of the project was led solely by Amy Priest (who has participated in the first two stages) as part of her continued work in the Yoder-Himes lab with the following goals: 1) examine the relationship of economic status and bacterial diversity at sampling sites; 2) obtain single colony *Ochrobactrum* sp. isolates (which predominated in the whole community analysis, appearing in 78 of the 87 communities), and determine their relative resistance to a spectrum of antibiotics; and 3) determine the ability of *Ochrobactrum* to transfer its antibiotic resistance trait to an antibiotic sensitive bacterium.

The **primary hypotheses** of this project are that natural water environments will contain a wide array of different antibiotic resistant bacterial species and that characteristics of sampling sites, such as the site's habitat (natural water, agricultural water, wastewater, or commercial drains), or the relative affluence of the area as measured by median household income will influence overall bacterial diversity. Finally, environmental *Ochrobactrum* species resistant to vancomycin and colistin, will be resistant to additional antibiotics.

Materials and Methods

Collecting Antibiotic Resistant Isolates

General and Selective Solid Culture Medium

Tryptic Soy Agar (TSA), a general nutrient medium containing 15 g casein enzymatic digest, 5 g soybean meal enzymatic digest, 5 g sodium chloride, and 15 g agar per 1L of medium, was purchased as a mixture (Ward's Scientific) and prepared for culture plates per the manufacturer's instructions. The mixture was autoclaved at 121 °C and 15 p.s.i. for 30 minutes to

sterilize it. The mixture was gradually cooled in a 55 °C water bath prior to pouring ~30 mL aliquots into petri dishes. When indicated, vancomycin (50 µg/mL) and/or colistin (32 µg/mL) were added after cooling and prior to pouring. Tryptic Soy Broth (TSB; Ward's Scientific), a liquid medium used for these studies, was also prepared and stored in 100 mL aliquots at room temperature.

Sample Collection

As part of Dr. Himes' and Dr. Yoder-Himes' BIOL 501 class (Fall 2016), a group of 21 undergraduate and post-baccalaureate students (including Amy Priest) collected antibiotic resistant bacteria in environmental water sources across the Louisville, KY metropolitan area. Water sources were classified into four categories: natural water sources, such as lakes and rivers; agricultural waters, such as puddles and soil samples on or near crop fields; wastewater sources, such as standing water and sediments in drainage ditches and storm drains; and drains and sinks in commercial establishments. Additionally, for each sample, pH and temperature were recorded, and at each site, weather information, GPS coordinates, and time and date of collection were recorded. Sites were sampled using a sterile cotton or by collecting water in sterile container, and sterile swabs were then used to inoculate plates containing: plain TSA and TSA + vancomycin + colistin (plain TSA plates will be referred to as "non-selective" communities, while plates with the selective pressure of antibiotics will be referred to as "selective" communities). Plates were incubated at 37 °C for 12–18 hours. Unique colonies were chosen and re-streaked three consecutive times to ensure isolation of pure single bacterial isolates (However, some colonies that initially grew were not able to be re-isolated). The remaining growth on the original nonselective and selective plates from each sampling site was collected using a sterile swab to wipe the surface

of the plate and preserved as whole community freezer stocks by swirling the inoculated swab in

TSB + 20% glycerol and then storing at –80 °C.

Preparing Antibiotic Resistant Samples for 16S rRNA Sequencing

Genomic DNA Extraction

Each antibiotic resistant single colony isolate was suspended in liquid culture of 5 mL TSB

and incubated overnight at 37 °C. Genomic DNA extraction was performed on each isolate using

the EtNA protocol [28]. Briefly, 100 μL of bacterial overnight culture and 455 μL of EtNA DNA

extraction reagent (5.5 mL of 2M NaOH, 35 mL of 96% ethanol, and 5 mL of 0.025M EDTA)

were mixed, heated for 10 minutes at 80 °C and spun for 10 minutes at 15,000 rpm in a microfuge.

The supernatant was removed and the pellet was resuspended in 100 μL of DNA resuspension

solution (5 mL of 0.5M Tris-HCl pH 8.0, 0.01 mL of 0.5M EDTA, pH 8.0, 0.5 mL of Triton X-

100, 0.25 mL of Tween 20, and 44.25 mL of sterile water).

Amplifying 16S rRNA using Polymerase Chain Reaction

The 16S ribosomal RNA gene sequences in bacteria differentiate between bacterial species

and thus is a means to identify unknown bacterial isolates [30–32]. The polymerase chain reaction

was used to amplify the 16S gene of 155 isolates by all the students in BIOL 501. (Note: multiple

trials of PCR occurred, due to initial failure of the amplification, and members of Dr. Himes' lab

also prepared isolates for sequencing). A master mix for 10 reactions (note: every student created

a Master Mix for their isolates but reagent amounts may have differed depending on the number

of isolates each student was preparing) was created for each sample, including: 250 μL NEBNext

High Fidelity 2× PCR Master Mix (BioLabs) DNA polymerase, 25 μL of 27F universal bacterial

primer (5'-AGAGTTTGATCMTGGCTCAG-3′), 25 μL of 1392R universal bacterial primer (5′-

ACGGGCGGTGTGTRC-3') [33], and 180 μL sterile water. The master mix was equally divided

in 10 PCR tubes and 2 μL of genomic DNA template was added to each tube. Additionally, positive

and negative controls were prepared for each reaction plate shared by multiple students. As a

negative control 2 μL of DNA resuspension solution was added to a tube of reagents, and as a

positive control 2 μL of *P. aeruginosa* strain PAO1 (a genome prep provided by Drs. Himes and

Yoder-Himes) was added to a tube of reagents. An Applied Biosystems (ABI) PCR thermal cycler

ThermoFisher) was used to run the NEBNext Protocol: 1) 98 °C for 5 minutes, 2) 35 cycles of 98

°C for 15 seconds, then 55 °C for 30 seconds, then 72 °C for 1 minute, 3) 72 °C for 10 minutes,

and 4) 4 °C hold. Gel electrophoresis verified the efficacy of PCR amplification through the

production of a ~1300 bp band. Samples with this size band were purified using the QIAQuick

PCR purification kit (Qiagen) per the manufacturer's instructions and eluted in 50 μL sterile water.

To assess the samples' nucleic acid quality, a NanoDrop microvolume spectrophotometer (Thermo

Scientific) was used. In general, pure nucleic acids have an A_{260}:A_{280} ratio between 1.8–2.2 and an

A_{260}:A_{230} ratio under 2 [29].

Identifying Antibiotic Resistant Isolates

Individual Isolate Sequencing and Database Analysis

Purified PCR amplicons were sent to University of Kentucky HealthCare Genomics Core

Laboratory for Sanger sequencing. The resultant DNA sequences were manually trimmed by each

student for quality using BioEdit Sequence Alignment Editor software [34]. Sequences were then

compared to two different databases: the National Center for Biotechnology Information's

GenBank non-redundant (nr) nucleotide database running the Basic Local Alignment Search Tool

(BLAST) algorithm [35] and Michigan State University's Ribosomal Database Project (RDP; a

repository database solely of small subunit rRNA sequences, including both 16S and 18S sequences) [36]. Using DNA sequence matches from the databases, samples were identified to the family, genus, or species (when possible) level. To characterize certainty of the designation, Genbank generates an e-value, which indicates the likelihood that two sequences are from different sources (the smaller the e-value, the less likely the two sequences are from different sources) while RDP generates a percent confidence score (the higher the confidence, the more likely the two sequences are from the same source).

Whole Community Analysis

Metagenome Sequencing

Drs. Yoder-Himes and Himes extracted genomic DNA from 96 bacterial communities collected during the BIOL 501 course. This corresponded to 48 paired samples grown on both TSA alone and TSA + vancomycin + colistin. Genomic DNA from each metagenomics sample was isolated using the protocol found in [37]. A small portion (10 μg) of each sample was submitted to the Kentucky Bioinformatics Research Informatics Network for sequencing. Libraries were constructed using Illumina's 16S library preparation guide and Illumina's Nextera Index Kit (FC-121-1012) and quantitated using the Qubit dsDNA BR Assay Kit in a Qubit 2.0 Fluorometer. Pooled libraries were sequenced using an Illumina MiSeq Reagents kit v3 (600 cycles) (MS-102-3003) on an Illumina MiSeq instrument.

QIIME (version 1.8), a bioinformatics pipeline, was used to analyze the samples [38]. Using the QWRAP (v. 2) pipeline, paired ends reads were merged and assigned phylogeny, based on the procedure previously described in [39] using Greengenes (version 4feb2001), a 16S rRNA gene database. QIIME assigned reads to taxonomic units—clusters of 16S rRNA sequences sorted

into groups based on similarity to reference taxonomy sequences [40]. Operational taxonomic units (OTUs) were defined as isolates sharing >97% similarity in their sequences which represent current definitions of bacterial species. For complete details on QWRAP protocols and commands used, see **Supplementary Protocol 1** in the Appendix.

Group Comparisons

Alpha diversity is a measure of diversity within a sample. QWRAP was used to examine alpha diversity through the Shannon diversity index and generate graphs and heatmaps to visualize alpha diversity. Beta diversity, a measure of diversity between samples, was estimated using principle coordinate analyses of weighted Unifrac distance and was visualized using Emperor software [41]. The QWRAP script_adv.sh function was used to create phylogenetic trees of the samples based on beta diversity estimates.

Analysis of Community Locations and Income Data

GPS coordinates were recorded for each sampling site and used to obtain physical addresses (if not previously recorded). With this information, median household income information was determined using an interactive data map at http://www.city-data.com/. The map presented income data (obtained from the United States Census Bureau) for many smaller sub-sections of within zip codes. The map was zoomed in to visualize streets and intersections to find the approximate location of the street address. Google Maps was used as a cross-check to ensure the locations matched. For comparison purposes, income data was also clustered into "high," "medium," and "low" categories. Household income percentiles for Louisville, KY were obtained from https://statisticalatlas.com/metro-area/Kentucky/Louisville/Overview. Household income

greater than the 60th percentile (>$55,700) of the median household income in Louisville was classified as "high." Household income between the 40th and 60th percentiles ($34,400-$55,700) was classified as "medium." Household income less than the 40th percentile (<$34,400) was classified as "low."

Income data was compared to each of the following metrics: alpha diversity (Shannon diversity index), genera and species classifications (both antibiotic and control), sample temperature, and sample pH). Using GraphPad Prism (version 5.0), linear regressions and correlations were calculated for each metric. Finally, with assistance from Dr. Sarah Emery, an analysis of covariance (ANCOVA) was conducted using Systat statistical software. ANCOVA determines the interaction and effects of an independent variable on a dependent variable, while considering additional variables that may vary along with the dependent variable (covariates). Median household income, habitat (agricultural water, natural water, drains, wastewater), sample pH, and sample temperature were set as independent co-variants to determine whether any had a significant effect on sample diversity using Shannon diversity index measures.

Ochrobactrum sp. Analysis

Phenotypic Differentiation of Isolates

Whole community metagenome data was used to determine the ten communities with the most occurrences of *Ochrobactrum* species. Based on literature searches regarding this genus [42], an isolation medium was made using: 35 g Columbia broth base, 15 g agar, and sterile water to 1L final volume. The following antibiotics were added to select for *Ochrobactrum*: colistin (32 µg/mL), vancomycin (50 µg/mL) and ceftazidime (30 µg/mL). Each unique colony was re-streaked three times to ensure a pure culture was obtained.

Gram-staining and catalase metabolism tests were used to remove isolates from consideration that were Gram-positive and/or catalase negative, as *Ochrobactrum* sp. are Gram-negative and catalase positive [43,44]. Gram-staining was done by preparing a bacterial emulsion in a drop of water on a glass slide. The emulsion was dried and heat-fixed by waving over a Bunsen burner. Crystal violet stain was applied to the slide for 1 minute and then rinsed with deionized water. Iodine was applied as a mordent for 30 seconds and rinsed with deionized water. Then, 70% ethanol was dripped over the slide, until the run-off was clear. Finally, safranin stain was applied to the slide for 1 minute and rinsed with deionized water. The slides were dried and viewed under a microscope. Cells that appeared purple are Gram-positive while cells that appeared red/pink are Gram-negative To conduct a catalase test, bacterial emulsions in a drop of water were prepared on a glass slide. Hydrogen peroxide was added to the emulsion. Bubbling of the mixture indicated metabolism of hydrogen peroxide and was a catalase positive result.

Genomic DNA Extraction and 16s rRNA Amplification

The Wizard Genomic DNA Purification Kit (Promega) was used to extract genomic DNA from each isolate. PCR was used to amplify the 16S gene with the following mix of reagents for each reaction (from Qiagen): 10 µL 5× PCR Buffer, 10 µL Q Buffer, 2 µL 27F universal bacterial primer, 2 µL 1392R, 1 µL HotStarTaq DNA polymerase, and 25 µL of sterile water. To the reaction mix, 2 µL of each sample was added, including water as a negative control and a genomic prep of a lab strain of *P. aeruginosa* as a positive control. The reaction was run under these conditions: 1) 95 °C for 5 minutes, 2) 35 cycles of 95 °C for 1 minute, then 55 °C for 1 minute, then 72 °C for 2 minutes, 4) 4 °C hold. Gel electrophoresis verified the presence of a band approximately 1300 bp in size. Samples were purified using the QIAQuick PCR purification kit (Qiagen), were assessed

for nucleic acid quality using previously described methods, and were sent to Macrogen in Baltimore, MD for sequencing.

Minimum Inhibitory Concentration (MIC) Assays

Antibiotic Preparations

Antibiotic stock solutions were made using the antibiotics and their respective solvents listed in **Table 1.** Concentrations of 26.5 mg/mL were obtained by measuring 128 mg of each antibiotic and suspending in 5 mL of solvent. To guarantee sterility, solutions with water as a solvent were filtered through a 0.20 μm syringe filter.

Preparation of 96-well Plates

Twelve 96-well plates were utilized (triplicate assays of each isolate were performed and only one type antibiotic was added per plate). A multi-channel pipette was used to load 196 μL of Mueller-Hinton (MH) broth (Sigma-Aldrich) into Row A of every plate. One hundred μL MH broth were loaded into Rows B-H. In Row A, 4 μL antibiotic solutions (25.6 mg/mL) were added to obtain a final concentration of 512 μg/mL and serially diluted 2-fold by pipetting 100 μL of the previous row to the subsequent row through Row G. After this process rows A-F contained 100 μL of solution while Row G contained 200 μL of solution. No antibiotics were added to Row H, which served as a positive control containing only 100 μL MH broth for strain growth.

Table 1. Antibiotics used in this study				
Antibiotic	Classification	Targets[†]	Mechanism	Solubility (mg/mL)
Vancomycin	Glycopeptides	G +, some G -	Disrupts cell wall formation by preventing linkage of NAG and NAM subunits in peptidoglycan layer	Water – 100
Colistin	Polymyxins	G -	Disrupts outer cell membrane, via binding to lipopolysaccharide	Water – 50
Tobramycin	Aminoglycosides	G -, some G +	Binds to 30S ribosome subunit, inhibiting protein synthesis	Water – 50
Carbenicillin	Penicillins	G - and G +	Prevents peptidoglycan cross-linkages, inhibiting cell wall synthesis	Water – 50
Kanamycin	Aminoglycosides	G -, some G+	Interferes with transcription initiation by binding with 16S rRNA	Water – 50
Trimethoprim	Sulfonamides	G - and G +	Dihydrofolate reductase inhibitor; removing precursor for pyrimidine synthesis	DMSO – 50
Tetracycline-HCl	Tetracyclines	G - and G +	Blocks amino-acyl tRNA synthetases, inhibiting protein synthesis	Water – 10
Imipenem	Carbapenems	G - and G +	Inactivates transpeptidases, inhibiting cell wall synthesis	DMSO – 50

[†]G+ indicates Gram positive organisms, G- indicates Gram negative organisms

Final antibiotic concentrations for each row were as follows: Row A (512 μg/mL);

Row B (256 μg/mL); Row C (128 μg/mL); Row D (64 μg/mL); Row E (32 μg/mL); Row F (16

μg/mL); Row G (8 μg/mL); Row H (0 μg/mL). Overnight bacterial cultures were suspended in 10

mL of LB and were used to add 10 μL of each culture in 3 subsequent columns for all Rows A-H.

Each plate was wrapped with one layer of parafilm around the clear lid, to prevent desiccation of

samples. Additionally, the plates were stored in a tub lined with fully dampened paper towels to

further aid in prevention of drying. Finally, the plates were rotated side-to-side at approximately

20 rpm for 18 hours at 37 °C on a plate rocker. Plates were visually analyzed for turbidity,

indicating growth in that well for MIC analysis. Each isolate was analyzed in triplicate over two

experiments resulting in six replicates. Additionally, a Tecan Infinite 200 microplate reader was

used to quantify growth in each well. The plate was placed in the reader without the lid, and the

machine measured optical density at 595 nm (O.D.$_{595}$), obtaining an average reading of 10 independent flashes of light through each well. Each isolate was analyzed in triplicate over two experiments resulting in six replicates.

To examine the samples' growth at lower antibiotic levels, the above process was repeated with the following concentrations: Row A (32 μg/mL); Row B (16 μg/mL); Row C (8 μg/mL); Row D (4 μg/mL); Row E (2 μg/mL); Row F (1 μg/mL); Row G (0.5 μg/mL); Row H (0 μg/mL). Once again, samples were rotated overnight at 20 rpm for 18 hours at 37℃ and read as described above. Each replicate reading was normalized as a percentage of the positive control (0 μg/mL) in Microsoft Excel. One-way ANOVA analyses using Tukey's multiple comparison post-test were performed in GraphPad Prism v5.0.

Assessing Ochrobactrum's Ability to Transfer Resistance

Plasmid Determination

The QIAprep Spin Miniprep Kit (Qiagen) was used per manufacturer's protocols to extract any plasmids from each strain. Spectrophotometry (NanoDrop) was used to examine the presence of DNA in the resultant samples.

Mating Assays

The transfer of antibiotic resistance was tested between one isolated *Ochrobactrum* sample and a freezer stock of *Pseudomonas aeruginosa* PAO1. Both samples were plated individually and together on plain LB plates to determine if they could be phenotypically distinguished (*Ochrobactrum* appeared as white colonies, which *P. aeruginosa* appeared noticeably different as yellow–green colonies). Additionally, the samples were plated on selective plates (LB +

vancomycin 50 µg/mL and LB + colistin 32 µg/mL) to confirm *Ochrobactrum*'s resistance to and *P. aeruginosa*'s susceptibility to both antibiotics.

To prepare mating assays, liquid cultures of each bacteria were grown overnight at 37 °C. As controls, 600 µL of each liquid culture were placed in separate microcentrifuge tubes. Additionally, 200 µL of *Ochrobactrum* (donor) liquid culture and 400 µL of *P. aeruginosa* (recipient) liquid culture were added to a mixed culture microcentrifuge tube. The tubes were spun at 1,500 rpm for 1 minute to concentrate the bacteria. The supernatant was removed and the pellet resuspended in 100 µL of plain LB. The entire resuspension mixture was pipetted on the center of an LB plate, allowed time to dry, and then incubated overnight at 37 °C.

The controls and mating mix were scraped off the plates and resuspended in 1 mL of plain LB. The tubes were diluted by a factor of 10^6. To do this dilution, 10 µL of culture was aliquoted in 1 mL of plain LB twice. Then 10 µL of each mix was pipetted and spread on both plain LB and selective plates (LB + 50 µg/mL vancomycin and LB + 32 µg/mL colistin). This process was conducted three times to confirm results.

Results

Freshwater environments may be a hub for the transfer of mobile genetic elements that can contain antibiotic resistance genes. Vancomycin, a glycopeptide antibiotic, and colistin, a polymyxin antibiotic, have previously been used as last resort antibiotics. Emerging clinical isolates show resistance to both vancomycin and colistin which is endangering the efficacy of antibiotics used in the clinical setting [19, 20, 23]. Learning more about these antibiotic resistant bacteria and the communities with which they interact is an imperative precursor for determining

how to combat the evolution and spread of antibiotic resistance. The first stage of our study aimed

to obtain information on local antibiotic resistance bacteria by studying the following four different

water-based habitats: agricultural waters, wastewaters, commercial drains, and natural water.

These water sources were chosen by the members of BIOL 501 based on their unique levels of

human interference, nutrient availability, and studies in the primary scientific literature.

Identifying Antibiotic Resistant Isolates

From September through October 2016, in total, 193 individual bacterial isolates were

collected by a team of undergraduate and post-bac students from the following habitats throughout

the Louisville regional area: wastewaters (50 isolates), agricultural waters (49 isolates), natural
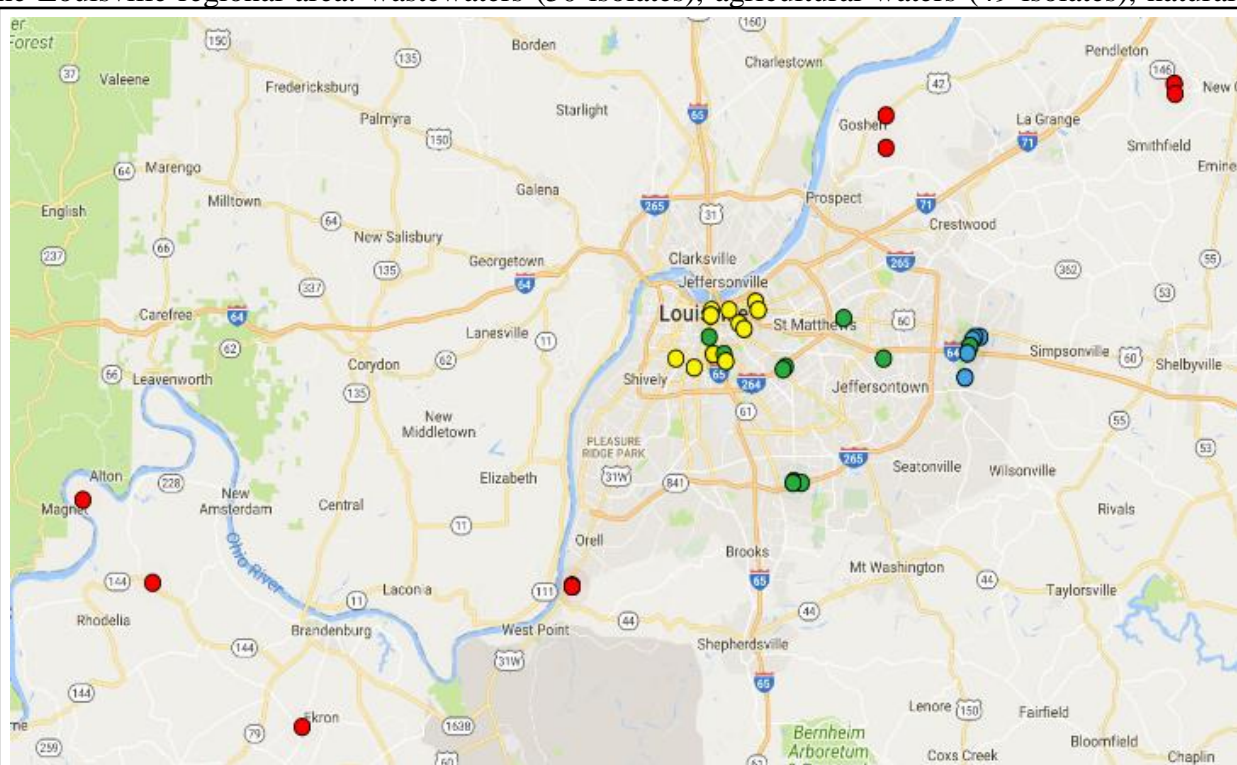
Figure 1. Sampling locations for this study within the greater Louisville area. Each dot represents a unique geographic location where sampling was conducted. The colors of the dots correspond to category of sample obtained from the location: blue – natural waterways; green – wastewaters; red – agricultural waters; yellow – commercial drain waters.

waters (48 isolates), commercial drains (46 isolates) (**Figure 1**). Genomic DNA from each isolate

24

was harvested by all members of BIOL 501 for their respective isolates and used as templates for

PCR to amplify the 16S rRNA gene which is used to identify bacteria at the species level [30–32].

Amy Priest isolated 11 of the wastewater samples; however, 3 were not identified in the initial trial

during the BIOL 501 class.

Of these 193 isolates obtained and prepared by members of BIOL 501 class and Dr. Himes

lab members, 155 showed amplification of the 16S rRNA gene with bands of the expected size.

These amplicons were purified and sequenced. Of the 155 samples sent for sequencing, sequences

from 98 samples allowed for classification of the isolate at least to the family taxonomic level

(**Table 2**). To identify the unknown isolates, the 16S sequences were compared to known strain

sequences in Genbank and RDP databases and used to assign a taxonomic designation for each

isolate.

Table 2. 16S rRNA Sequencing Genbank and RDP Designations						
ID	**Category**	**Length (bp)**	**Nucleotide Genbank Designation**	**BLAST e-value**	**RDP Classification**	**RDP value**
GRS2	Wastewater	410	*Pseudochrobactrum* sp.	0.0	*Pseudochrobactrum* sp.	100%
WMB W1	Wastewater	440	*Sphingobacterium* sp.	1.00E-134	*Sphingobacterium* sp.	60%
WMB W2	Wastewater	608	*Roseomonas* sp.	0.0	*Roseomonas* sp.	100%
**KG S2	Wastewater	587	*Sphingobacterium* sp.	7.00E-178	*Sphingobacteriaceae* gen. sp.	62%
**KG S8	Wastewater	547	*Roseomonas* sp.	1.00E-125	*Acetobacteraceae* gen. sp.	98%
**CM G1	Wastewater	370	*Sphingobacterium* sp.	3.00E-165	*Sphingobacterium* sp.	95%
**CM G2	Wastewater	346	*Chryseobacterium* sp.	2.00E-141	*Flavobacteriaceae* gen. sp.	97%
**W MS3	Wastewater	370	*Pseudochrobactrum* sp.	2.00E-166	*Pseudochrobactrum* sp.	95%
**GR W1-2	Wastewater	420	*Serratia* sp.	0.0	*Enterobactericeae* gen. sp.	100%
**W MW1	Wastewater	370	*Ochrobactrum* sp.	0.0	*Ochrobactrum* sp.	85%

**W MW2	Wastewater	325	*Serratia* sp.	2.00E-126	*Enterobacericeae gen. sp.*	98%
*6M M1	Wastewater	358	Uncultured Bacterium	1.00E-118	*Epilithonimonas* sp.	73%
*SEW 2	Wastewater	395	*Serratia marcescens*	2.00E-101	*Mangrovibacter sp.*	41%
3OLA RGE	Wastewater	631	*Serratia marcescens*	0.0	*Serratia* sp.	64%
3OMS	Wastewater	628	*Serratia marcescens*	0.0	*Serratia* sp.	68%
GRS1	Wastewater	412	*Serratia* sp.	0.0	*Serratia* sp.	99%
CMD 1	Wastewater	640	*Serratia marcescens*	0.0	*Serratia* sp.	100%
OXC M1	Wastewater	671	*Chryseobacterium jujuense*	0.0	*Chryseobacterium* sp.	100%
OXC M2	Wastewater	500	*Serratia* sp.	0.0	*Serratia* sp.	80%
OXC M3	Wastewater	401	*Providencia* sp.	0.0	*Providencia* sp.	100%
MSD MLSS	Wastewater	372	*Flavobacterium* sp.	0.0	*Flavobacteriaceae gen. sp.*	87%
*MSD RAS	Wastewater	435	*Serratia marcescens*	0.0	*Flavobacteriiaceae gen. sp.*	98%
LFW2	Wastewater	555	*Serratia marcescens*	0.0	*Enterobacericeae gen. sp.*	100%
TDW 2	Wastewater	476	*Providencia vermicola*	0.0	*Providencia* sp.	100%
TDW 1	Wastewater	1251	*Aeromonas hydrophila*	0.0	*Aeromonas* sp.	99%
TDS3	Wastewater	579	*Serratia marcescens*	0.0	*Serratia* sp.	100%
TDS2	Wastewater	359	*Klebsiella variicola*	0.0	*Enterobacericeae gen. sp.*	100%
LFM UD	Wastewater	570	*Serratia marcescens*	0.0	*Enterobacericeae gen. sp.*	100%
Dr37	Drains	592	*Sphingobacterium multivorum*	0.0	*Sphingobacterium* sp.	100%
Dr38	Drains	219	*Serratia fonticola*	0.0	*Enterobacericeae gen. sp.*	100%
Dr39	Drains	591	*Pseudomonas fuscovaginae*	0.0	*Pseudomonadaceae gen. sp.*	100%
Dr40	Drains	627	*Pantoea ananatis*	0.0	*Enterobacericeae gen. sp.*	100%
Dr42	Drains	435	*Serratia fonticola*	0.0	*Serratia* sp.	100%
Dr44	Drains	506	*Serratia marcescens*	0.0	*Serratia* sp.	100%
Dr46	Drains	540	*Providencia rettgeria*	0.0	*Providencia* sp.	100%
*MR2 6	Drains	558	Bacterium strain	0.0	*Delftia* sp.	100%
MR28	Drains	622	*Pseudochrobactrum sp.*	0.0	*Pseudochrobactrum sp.*	100%

*MR29	Drains	191	Bacterium strain	4.00E-92	*Delftia* sp.	100%
*MR30	Drains	430	Bacterium strain	0.0	*Sphingobacteriaceae* gen. sp.	100%
MR31	Drains	561	*Pseudochrobactrum asaccharolyticum*	0.0	*Pseudochrobactrum* sp.	100%
MR34	Drains	516	*Brucellaceae* gen. sp.	0.0	*Pseudochrobactrum* sp.	100%
CSAB 6-1-1	Drains	573	*Serratia marcescens*	0.0	*Serratia* sp.	100%
CSAB 6-2-1	Drains	473	*Stenotrophomonas maltophilia*	0.0	*Stenotrophomonas* sp.	100%
CSAB 6-2-2	Drains	502	*Stenotrophomonas maltophilia*	0.0	*Stenotrophomonas* sp.	100%
CSAB 6-2-3	Drains	471	*Enterobacter cloacae*	0.0	*Enterobactericeae* gen. sp.	100%
*W1Shade	Natural Waters	370	*Aeromonas hydrophila*	1.00E-114	*Paraferrimonas* sp.	22%
*I2A	Natural Waters	300	*Serratia marcescens*	1.00E-132	*Samsonia* sp.	40%
*C4A	Natural Waters	410	*Rahnella aquatillis*	0.0	*Ewingella* sp.	16%
C2A	Natural Waters	350	*Serratia marcescens*	9.00E-165	*Serratia* sp.	83%
*C1C4	Natural Waters	270	*Serratia marcescens*	1.00E-169	*Rahnella* sp.	53%
I1A2	Natural Waters	460	*Aeromonas hydrophila*	0.0	*Aeromonas* sp.	80%
FF-RM1-1	Natural Waters	456	*Pectobacterium carotavorum*	0.0	*Pectobacterium* sp.	82%
FF-RM1-2	Natural Waters	373	*Aeromonas* sp.	2.00E-177	*Aeromonas* sp.	95%
*SH-RM2-2	Natural Waters	439	*Proteus mirabilis*	2.00E-73	*Enterobactericeae moellerella*	17%
*SH-RM2-3	Natural Waters	520	*Flavobacterium* sp.	2.00E-48	*Petrimonas* sp.	14%
SH-RM2-4	Natural Waters	550	*Proteus mirabilis*	0.0	*Proteus* sp.	76%
SH-RM2-5	Natural Waters	298	*Providencia stuartii*	5.00E-36	*Providencia* sp.	12%
SH-RB2-1	Natural Waters	410	*Enterobacter* sp.	1.00E-64	*Enterobacteriaceae Moellerella*	22%
SH-RB2-2	Natural Waters	480	*Proteus mirabilis*	3.00E-92	*Enterobacteriaceae* gen. sp.	97%

*SH-RB2-3	Natural Waters	380	*Proteus mirabilis*	5.00E-108	*Cosenzaea* sp.	59%
*SH-RB2-4	Natural Waters	170	*Proteus mirabilis*	1.00E-104	*Cosenzaea* sp.	21%
*SH-SM1-2	Natural Waters	360	*Serratia* sp.	8.00E-131	*Cosenzaea* sp.	40%
SH-SM2-1	Natural Waters	541	*Proteus* sp.	0.0	*Proteus* sp.	43%
SH-SM2-2	Natural Waters	551	*Proteus* sp.	0.0	*Proteus* sp.	100%
SH-SM2-3	Natural Waters	501	*Proteus vulgaris*	0.0	*Proteus* sp.	96%
SH-FPA1-1	Natural Waters	561	*Proteus mirabilis*	0.0	*Proteus* sp.	100%
SH-FPA1-4	Natural Waters	629	*Proteus mirabilis*	0.0	*Proteus* sp.	80%
ML4A2	Natural Waters	525	*Aeromonas* sp.	0.0	*Aeromonas* sp.	80%
ML1A	Natural Waters	500	*Proteus mirabilis*	0.0	*Proteus* sp.	80%
*CHICK1B	Natural Waters	339	*Serratia marcescens*	9.00E-1.65	*Enterobacter* sp.	80%
I3A	Natural Waters	597	*Serratia marcescens*	0.0	*Serratia marcescens*	80%
*EEF1-C3	Agricultural Waters	400	*Pantoea ananatis*	6.00E-108	Unclassified	100%
MF2-O3	Agricultural Waters	410	*Chryseobacterium* sp.	0.0	*Chryseobacterium* sp.	100%
SF1-Y3	Agricultural Waters	445	*Providencia sneebia*	0.0	*Providencia* sp.	100%
PJF3-C3	Agricultural Waters	401	*Serratia marcescens*	0.0	*Serratia*	100%
DRAR-y1	Agricultural Waters	360	*Sphingobacterium sp.*	7.00E-162	*Sphingobacterium* sp.	100%
SDDR #1-p1	Agricultural Waters	483	*Serratia fonticola*	0.0	*Serratia* sp.	100%
SDDR #2-y2	Agricultural Waters	465	*Serratia fonticola*	0.0	*Serratia* sp.	100%
SFR# 1-y1	Agricultural Waters	362	*Yokenella sp.*	0.0	*Yokenella* sp.	40%
SFR# 3-r1	Agricultural Waters	462	*Serratia marcescens*	0.0	*Serratia* sp.	100%
2OC-1	Agricultural Waters	341	*Herbaspirillum frisingense*	3.00E-175	*Herbaspirillum* sp.	99%

GHF1-3	Agricultural Waters	469	*Pantoea ananatis*	0.0	*Pantoea* sp.	100%
GHF2-1	Agricultural Waters	300	*Pantoea ananatis*	3.00E-154	*Pantoea* sp.	98%
GHF2-2	Agricultural Waters	401	*Pantoea ananatis*	0.0	*Pantoea*sp.	100%
OT1-3	Agricultural Waters	297	*Sphingobacterium cladoniae*	2.00E-127	*Sphingobacterium* sp.	67%
*RI1-2	Agricultural Waters	141	*Serratia aquatillis*	3.00E-61	*Erwinia* sp.	51%
RI1-4	Agricultural Waters	320	*Serratia aquatillis*	6.00E-162	*Serratia*sp.	73%
DRAR-g1	Agricultural Waters	484	*Pantoea ananatis*	0	*Pantoea*sp.	100%
SFR-4-b1	Agricultural Waters	288	*Pseudochrobactrum asaccharolyticum*	3.00E-140	*Pseudochrobactrum* sp.	100%
PJF3-C3	Agricultural Waters	435	*Chitinophaga sp.*	0.0	*Chitinophaga* sp.	100%
SFR1-b2	Agricultural Waters	239	*Serratia marcescens*	2.00E-120	*Enterobacteriaceae* sp.	100%
SFS1-wf2	Agricultural Waters	458	*Pedobacter steynii*	0.0	*Pedobacter* sp.	100%
WP1+2.3	Agricultural Waters	337	*Chromobacterium* sp.	1.00E-174	*Chromobacterium* sp.	100%
WP1+3.3	Agricultural Waters	174	*Chromobacterium aquaticum*	4.00E-81	*Chromobacterium* sp.	97%
WC1+2.3	Agricultural Waters	386	*Providencia alcalifaciens*	0.0	*Providencia* sp.	100%
WP1+1.3	Agricultural Waters	279	*Chromobacterium aquaticum*	1.00E-138	*Chromobacterium* sp.	100%
WHF-1.3	Agricultural Waters	289	*Sphingobacterium sp.*	4.00E-143	*Sphingobacterium* sp.	100%
WHF+1.3	Agricultural Waters	149	*Providencia alcalifaciens*	1.00E-70	*Providencia* sp.	100%

*Denotes the 17 samples with discrepancies between BLAST and RDP designations
**Denotes the 8 samples isolated and identified by Amy Priest

Bacterial Isolates Taxonomic Designations

Of the 98 samples classified, 81 samples had matching designations between Genbank and RDP, with e-values near 0 in BLAST and RDP confidence levels (CL) greater than 60%. The frequencies of different taxonomic units were counted at both the genus and species level. Sixteen different genera were isolated, but the most commonly occurring genera (**Figure 2**) were

Serratia sp. (25 isolates), *Proteus* sp. (eight isolates), and *Sphingobacterium* sp. and *Providencia* sp. (seven isolates each). The most commonly occurring species were *Serratia marcescens* (14 isolates), *Proteus mirabilis* and *Pantoea ananatis* (five isolates each), and *Serratia fonticola* (four isolates) (**Figure 3**).
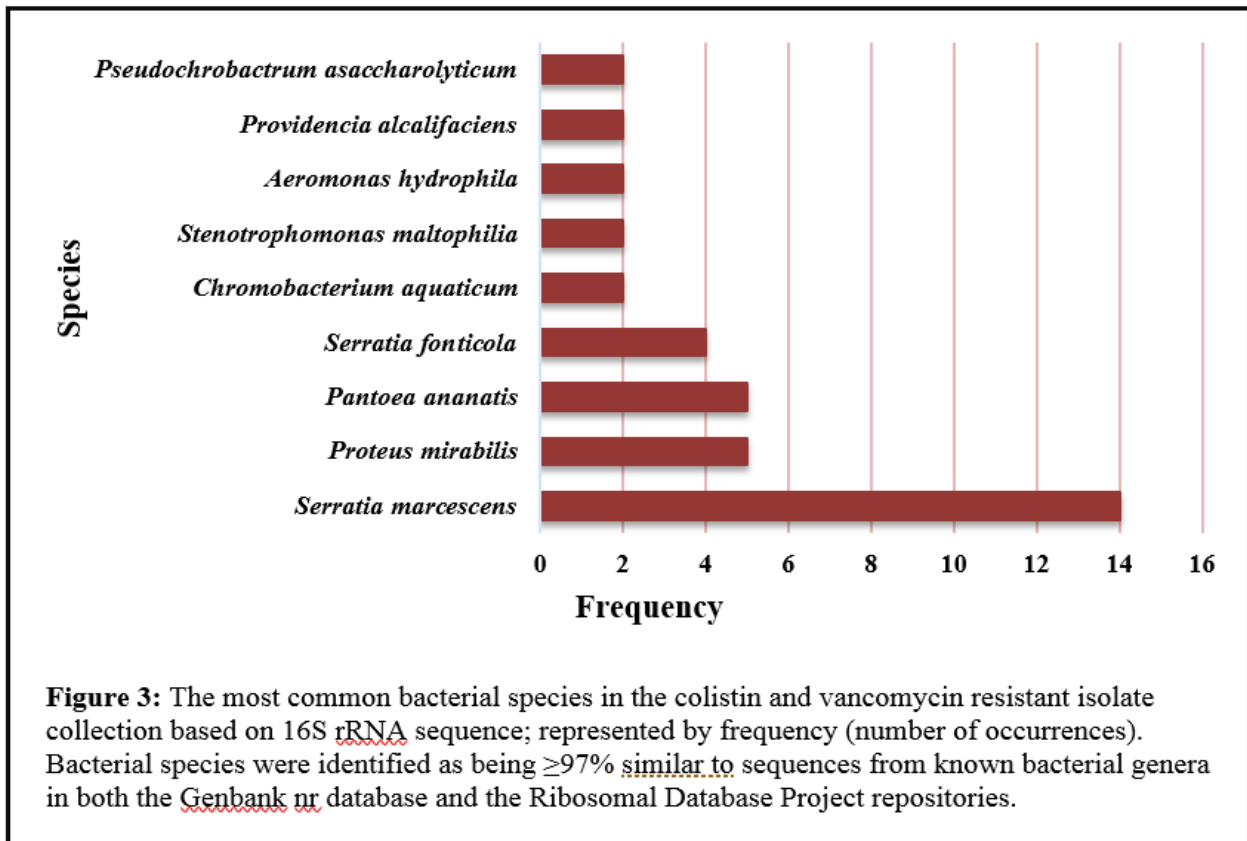

Figure 2. Bacterial genera in the colistin and vancomycin resistant isolate collection based on 16S rRNA sequence; represented by frequency (number of occurrences). Bacterial genera were identified as being ≥95% similar to sequences from known bacterial genera in both the Genbank nr database and the Ribosomal Database Project repositories.
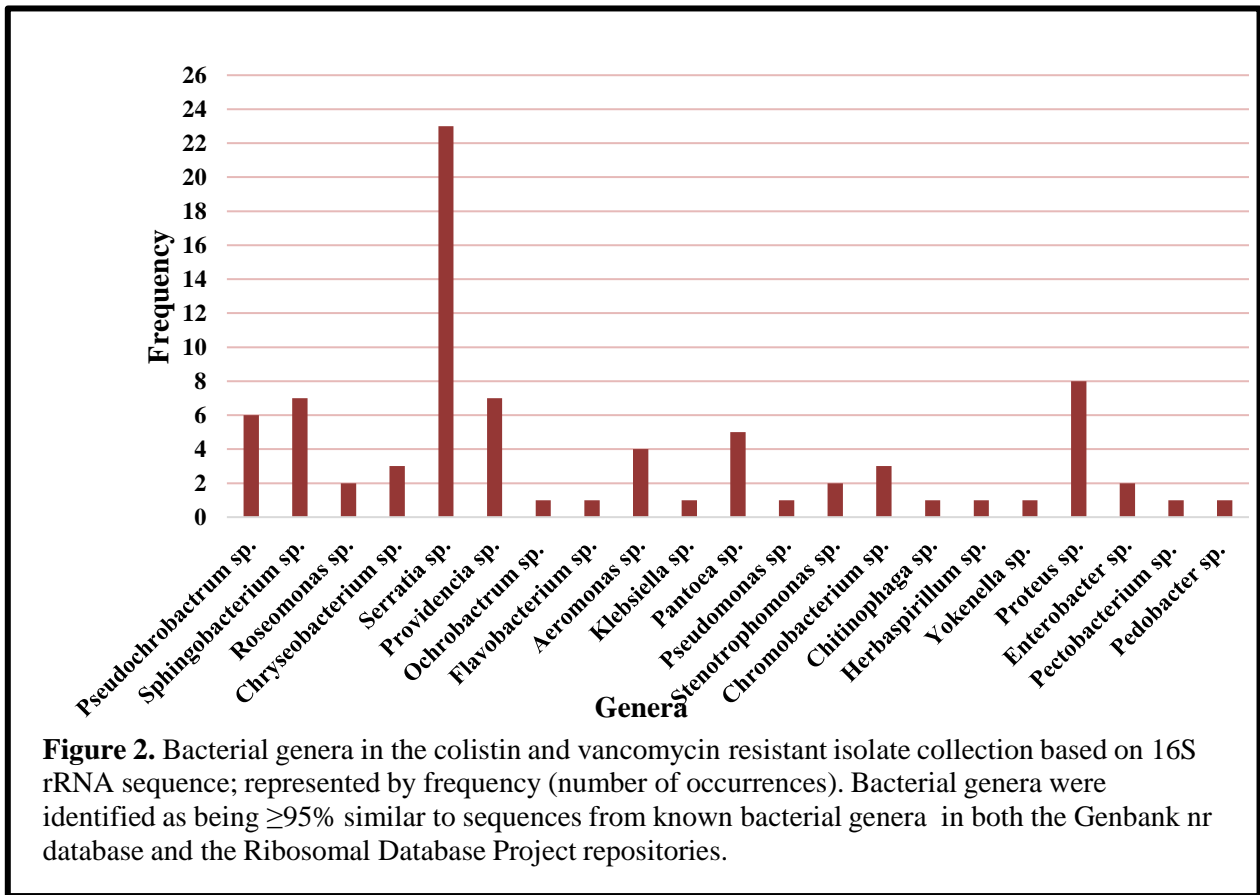
The remaining 17 samples showed discrepancies between Genbank and RDP classifications and/or low confidence levels, and thus their designations are less sure. However, of interest, two isolates (MR29 and MR30) were only classified as a "bacterium strain" in BLAST but in RDP both samples were classified as *Delftia* sp. (CL=100%). Additionally, 13 of the 98 samples were identified based on sequences less than 300 bp which may compromise the accuracy of those isolates designations.

From these analyses, it can be concluded that environmental waters can provide a habitat for the growth of a variety of bacterial species resistant to vancomycin and colistin. However, studying only single colony isolates has a few limitations: 1) a sampling bias may have occurred as students may have selected more colorful colonies for isolation. Additionally, some bacteria, particularly those sharing similar physical appearances or from densely populated communities may have been overlooked during single colony isolations; and 2) individual isolates do not provide detailed information about the communities as whole; for instance, whether certain habitats yield more diversity of antibiotic resistant bacteria than others.

Figure 3: The most common bacterial species in the colistin and vancomycin resistant isolate collection based on 16S rRNA sequence; represented by frequency (number of occurrences). Bacterial species were identified as being ≥97% similar to sequences from known bacterial genera in both the Genbank nr database and the Ribosomal Database Project repositories.

Whole Community Analysis

To expand on the data obtained from single colony isolates, whole community analyses (using preserved communities obtained while collecting individual isolates) were conducted. The goal of these analyses was to determine the bacterial composition of the entire community in order to provide sufficient information to make comparisons between selective and non-selective communities and between different habitats. Amy Priest contributed in group efforts to analyze sequences and was responsible for analyzing bacterial genera and species occurring in communities.

Obtaining Communities

Community samples (taken by the BIOL 501 class in conjunction with individual isolates) were collected from 34 individual geographical sites. Community samples were plated on both selective and non-selective media, allowed sufficient time to grow, and then were harvested by scraping the surface of the places and collecting in TSB + 20% glycerol. These communities were stored in the –80 °C freezer. Multiple samples were often taken at each location, resulting in a total of 157 unique communities collected. Of these 157 communities, 96 paired communities (48 communities grown on non-selective TSA medium and 48 communities grown under selection on TSA medium containing vancomycin and colistin from the same sites) were selected for 16S metagenomic sequencing to assess the overall diversity at these sites and the composition of antibiotic resistant species in overall microbial communities.

During sample analysis, nine community libraries failed to meet quality thresholds and were eliminated. After discarding those samples, 87 communities remained, including 39 paired selective and non-selective communities (In total, 45 communities grown on a selective antibiotic

medium remained and 42 communities grown on a non-selective control medium remained). From these communities, 6,079,485 reads were sequenced with an average of 69,309 sequences per sample (median = 62,814). These reads were used to determine the diversity and composition of each sample using QIIME, a free software used commonly in 16S metagenome studies [38]. These reads were assigned to taxonomic classifications in QIIME to a level with the highest confidence. For example, some reads could only be assigned to the phylum level, others could be assigned to the species level, and yet others could not even be assigned to the kingdom of bacteria.

Analysis of all the communities resulted in 215 unique taxonomic groupings. Some isolates present in the communities could not be determined at the genus level and were sorted only into higher taxonomic levels. Additionally, some bacteria could not be classified to any current taxonomic level and were sorted to an "unclassified" group. There were only 54 sequences categorized into this unclassified category.

Diversity within the Community Samples at the Genus Level

Of the 215 taxonomic groups, 116 of these groups were classified to the genus level; thus across all communities at least 116 different genera were present (**Figure 4**). The genus that was found in the greatest number of samples was *Ochrobactrum*, appearing in 78 samples out of the 87 final communities. In addition, the other most common genera across samples were: *Microbacterium* (31 samples), *Brevundimonas* (29 samples), *Serratia* (24 samples), and *Pseudomonas* and *Agrobacterium* (21 samples each) (**Figure 5**). Forty-one genera (35.34% of all genera classifications) appeared only one time, suggesting these genera may be unique to the habitat and rare across habitats.
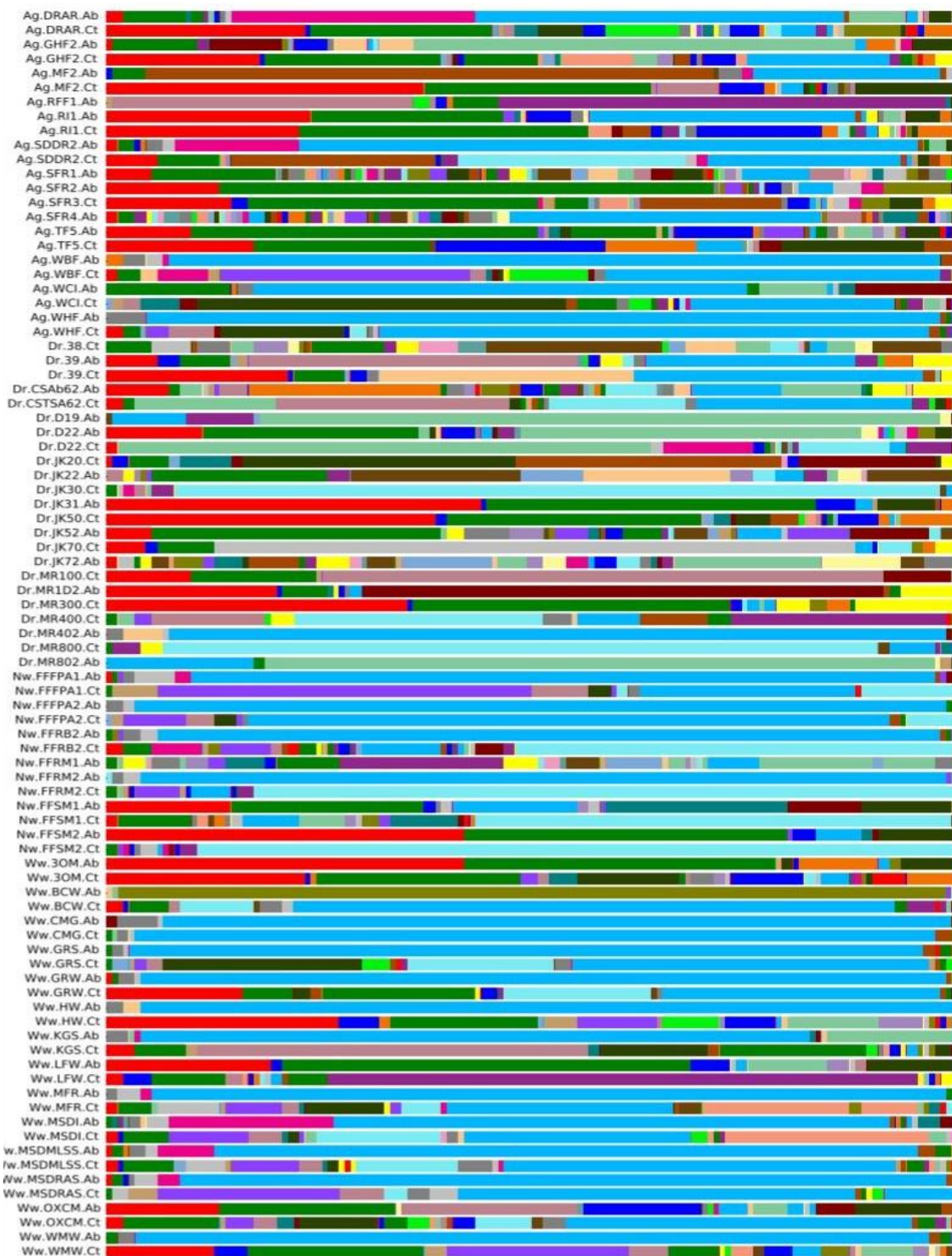
Selective and Non-Selective Community Genera Comparisons

Selective and non-selective paired communities were first analyzed by comparing how many OTUs occurred in both selective and non-selective paired communities and how many OTUs occurred in just one paired community (either selective or non-selective). Approximately 3.66% of OTUs occurred only in a selective community and not in the cognate non-selective community. More frequently, an OTU appeared in both paired communities or only in the non-selective community (8.01% of all OTUs). These results indicate that it is rare for a bacterium to be found only in a selective community and not in the cognate non-selective community.
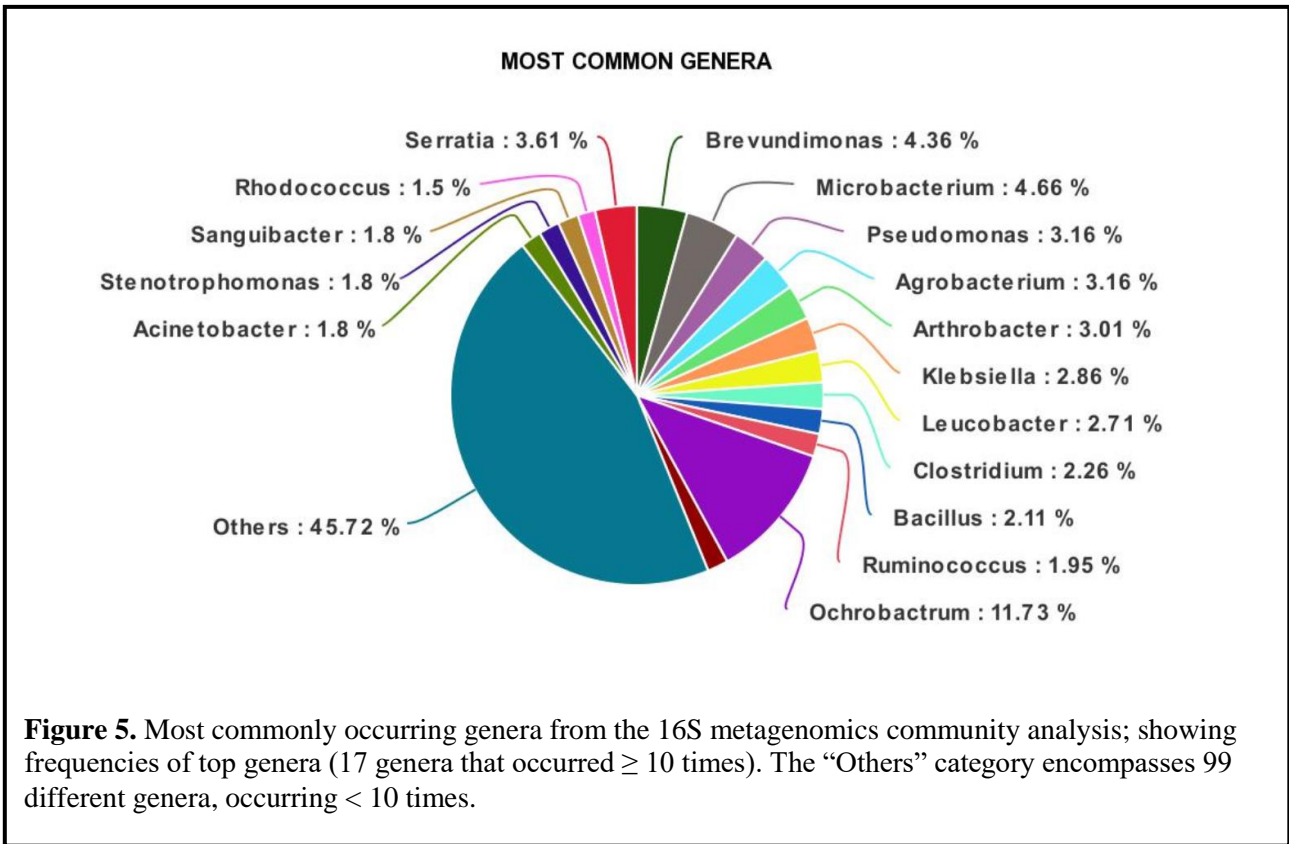
Quantities of different genera across all selective and non-selective communities were also analyzed. Of the 116 genera from all communities, there were 86 genera that appeared in selective communities and 79 genera that appeared in non-selective communities. There were 37 genera that only appeared in selective communities and 30 genera that only appeared in non-selective communities. This result was unexpected, as it was hypothesized that more bacteria would be able to grow under non-selective conditions. It is possible that less selective pressure allowed a non-resistant bacterium to dominate, crowding out other bacteria in non-selective communities.

Community Species Classifications

The diversity of microbes was then analyzed at the species level rather than the genus level to determine if any major differences could be observed. Analysis of the communities at the species taxonomic level resulted in 285 different taxonomic groups (**Figure 6**). Of the 285 OTUs, 70 were classified as species (24.56% of the sequences were sorted into species). This

Figure 4. 16S metagenomics community genera analysis from selective and non-selective samples. The left "Y" axis shows each community designation names. Ag indicates samples taken from agricultural sites, Dr from drains, Nw from natural waterways, and Ww from wastewaters. Following the community names, each sample names either an Ab designation indicating a sample grown in the presence of antibiotics or Ct indicating a sample grown under non-selective conditions. The colors in each row correspond to each taxonomic classification found in that sample.

Figure 5. Most commonly occurring genera from the 16S metagenomics community analysis; showing frequencies of top genera (17 genera that occurred ≥ 10 times). The "Others" category encompasses 99 different genera, occurring < 10 times.

low percentage indicates that species within many genera were not readily discernible by short read 16S sequencing. For example, *Ochrobactrum* dominated at the genus level as it was found in 78 samples, sometimes in very high abundance. However, at the species level, only six samples contained OTUs that could be classified into *Ochrobactrum* species (*O. intermedium* – three samples and *O. gallinifaecis* – three samples).

The frequencies of the top genera were compared between selective and non-selective communities. *Serratia* sp. and *Ochrobactrum* sp. appeared more frequently in selective communities than non-selective communities. *Serratia* sp. appeared in 15 selective and nine non-selective communities, while *Ochrobactrum* sp. appeared in 41 selective and 37 non-selective communities. Conversely, *Pseudomonas* sp., *Agrobacterium* sp., *Brevundimonas* sp., and *Microbacterium* sp. appeared more frequently in non-selective communities than selective

36

communities. *Pseudomonas* sp. appeared in 16 non-selective and five selective communities, *Agrobacterium* sp. appeared in 14 non-selective and seven selective communities, *Brevundimonas* sp. appeared in 24 non-selective and five selective communities, and finally, *Microbacterium* sp. appeared in 24 selective and seven selective communities.

The most commonly occurring species was *Brevundimonas diminuta*, appearing in 29 samples. In addition, other commonly appearing species were *Serratia marcescens* (24 samples), *Ruminococcus gnavus* (13 samples), and *Akkermansia muciniphila* and *Bacillus cereus* (10 samples each) (**Figure 7**). Thirty species only appeared one time, again indicating that these species are relatively rare across samples and/or habitats.
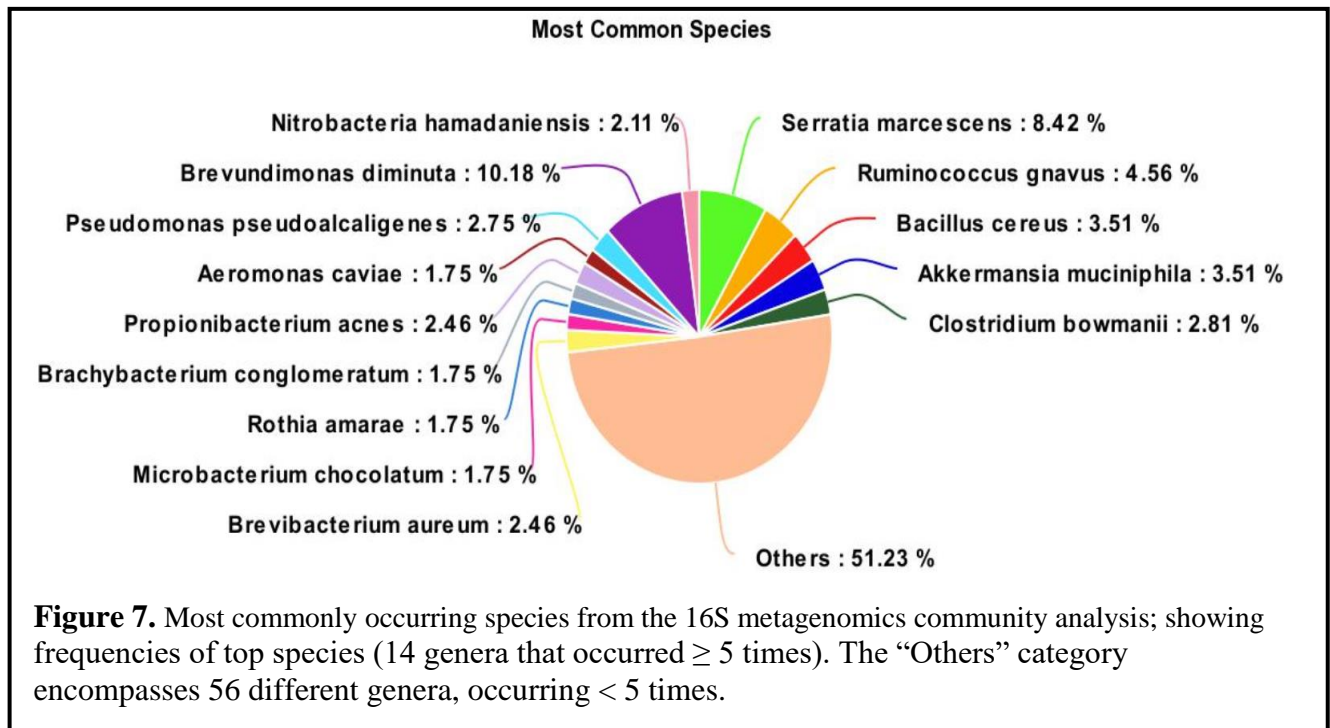
Selective and Non-Selective Community Species Comparisons

Of the 70 OTUs classified to the species level, there were 42 species that appeared in selective communities and 54 species that appeared in non-selective communities. There were 16 species that only appeared in selective communities and 27 species that only appeared in non-selective communities. This result was expected as antibiotics add selective pressures that may limit the growth of some bacteria.

The frequencies of the top species were compared between selective and non-selective communities. *R. gnavus, S. marcescens*, and *A. muciniphila* appeared more frequently on selective communities than non-selective communities. *R. gnavus* appeared in 8 selective communities and 5 non-selective communities, *S. marcescens* appeared in 15 selective communities and 9 non-selective communities, and finally, *A. muciniphila* appeared in 7 selective communities and 3 non-selective communities. Conversely, *B. diminuta* and *B. cereus* appeared more frequently on non-selective communities than selective communities. *B. diminuta*
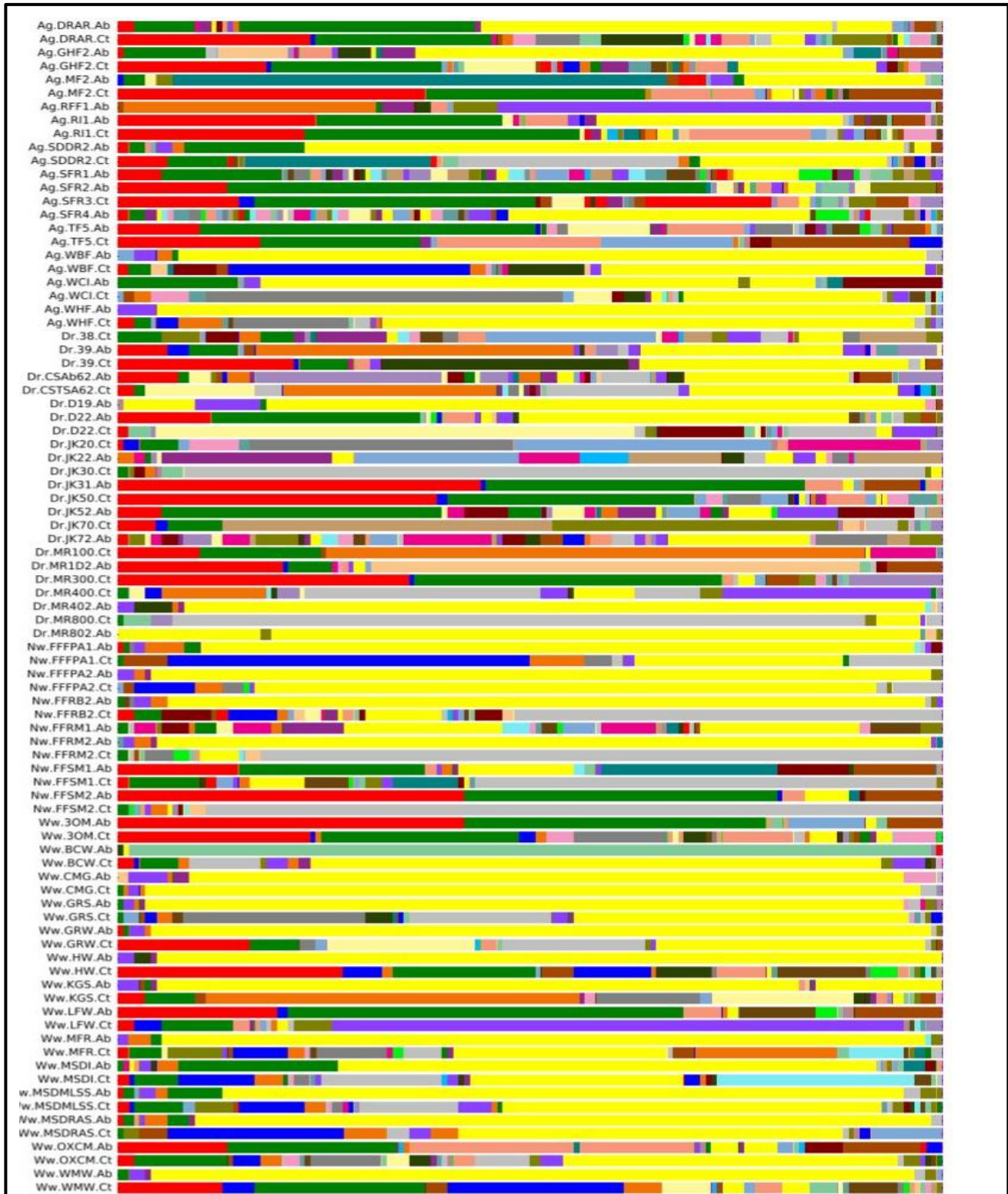
Figure 6. Community species analysis. The left "Y" axis shows each community designation. See Figure 4 for a description of sample names. The colors in each row correspond to each taxonomic classification found in that sample.

Figure 7. Most commonly occurring species from the 16S metagenomics community analysis; showing frequencies of top species (14 genera that occurred ≥ 5 times). The "Others" category encompasses 56 different genera, occurring < 5 times.

appeared in 24 non-selective communities and 5 selective communities, while *B. cereus* appeared in 8 non-selective communities and 2 selective communities. In general, these results indicate that perhaps some bacteria may grow more under selective conditions than non-selective conditions, while other bacteria may demonstrate the opposite trend in growth.
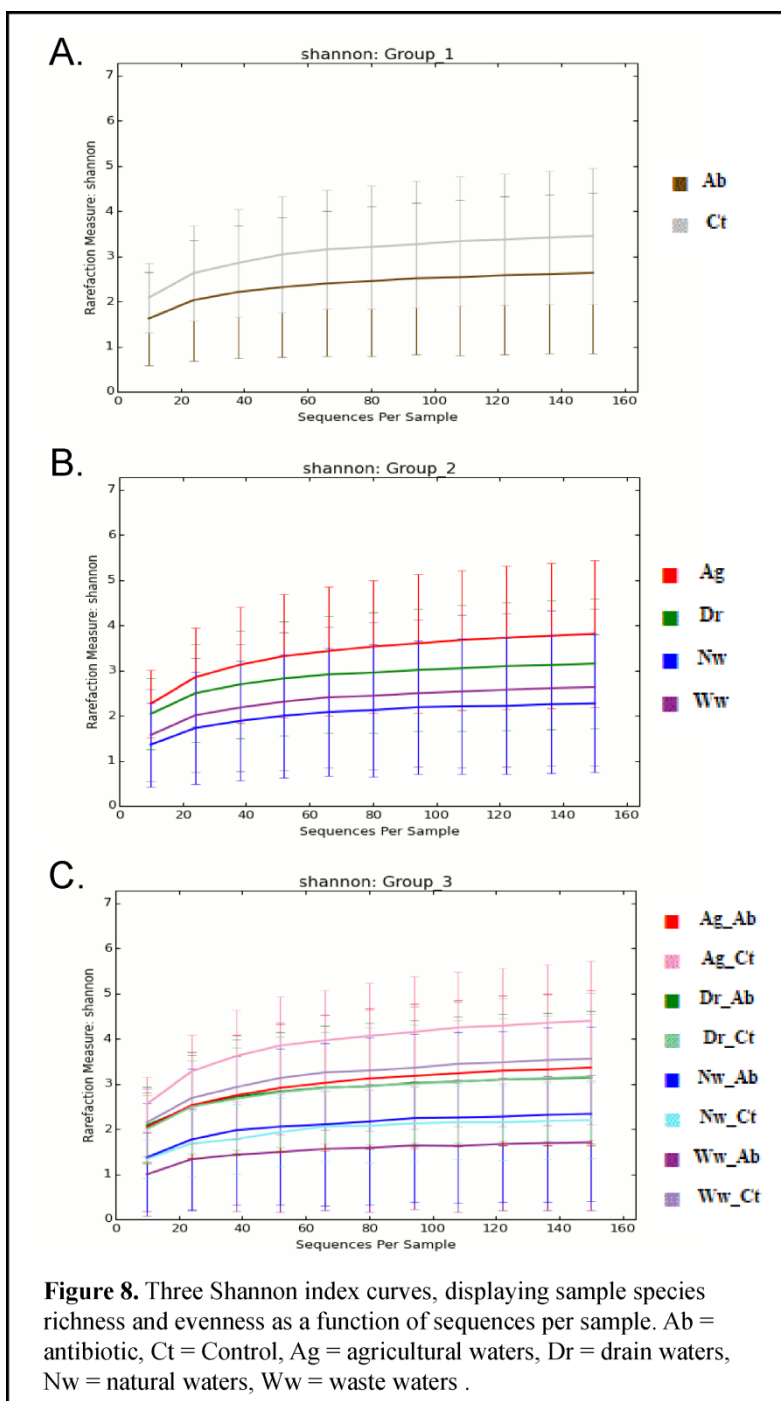
Group Comparisons: Alpha Diversity

Alpha diversity is a measure of species richness or the relative abundances of species within a given community [45]. One common metric, the Shannon diversity index (sometimes called the Shannon-Weiner index), weighs species evenness throughout a community [45]. The Shannon diversity metric describes the uncertainty of predicting the identity of the next individual encountered in the community. With greater species variability (greater richness), it will be more

challenging to predict the identity of the theoretical next species encountered—such communities will have higher Shannon diversity indices [45].

Alpha diversity tables were constructed in QIIME software to analyze species diversity within each community and rarefaction curves were plotted. Rarified curves adjust for any possible sampling bias present by randomly collecting a certain number of reads from a sample [50]. For example, one sample may have 1000 reads, while another may only have 500, by sheer quantity of reads, the samples with more reads could have more species present. Thus, sampling equal subsets of reads in each sample allows for fair comparison between samples.

The first analysis was conducted at a broad level by clustering all antibiotic samples together in a group and clustering all control communities together in a group. The broad-spectrum analysis

Figure 8. Three Shannon index curves, displaying sample species richness and evenness as a function of sequences per sample. Ab = antibiotic, Ct = Control, Ag = agricultural waters, Dr = drain waters, Nw = natural waters, Ww = waste waters .

40

revealed that the non-selective samples had greater measures of richness and evenness than the selective samples (**Figure 8A**). To test if the groups were significantly different, non-parametric t-tests were conducted and non-selective and selective were significantly different (p = 0.028). This result was expected due to the selective nature of antibiotics. In both groups, as the number of sequences in each sample increased, diversity and evenness briefly increased before reaching a plateau, indicating that most OTUs in each sample were accounted for in the data.

Subsequent analyses arranged the communities by habitat, or then further divided each habitat into antibiotic and control communities. Rarefaction curves were also generated according to habitat: drains, natural waters, wastewaters, and agricultural waters (**Figure 8B**). The agricultural samples showed the greatest amount of species richness and evenness as measured by the Shannon diversity index. Species diversity and richness decreased sequentially in drain, wastewater, and natural water samples, respectively. However, only agricultural waters and natural waters groups were significantly different (p = 0.042).

Finally, habitats were further divided into selective and non-selective communities, resulting in eight different categories. The non-selective agriculture samples had the greatest species richness and evenness while the non-selective wastewater samples showed the least species richness and evenness (**Figure 8C**). In agricultural waters and wastewaters, the non-selective samples had a much greater diversity compared to their associated selective samples. However, in natural water and drain samples, the Shannon diversity indices were very similar between selective and non-selective samples. However, only non-selective agricultural waters and selective wastewaters groups were significantly different (p = 0.028).
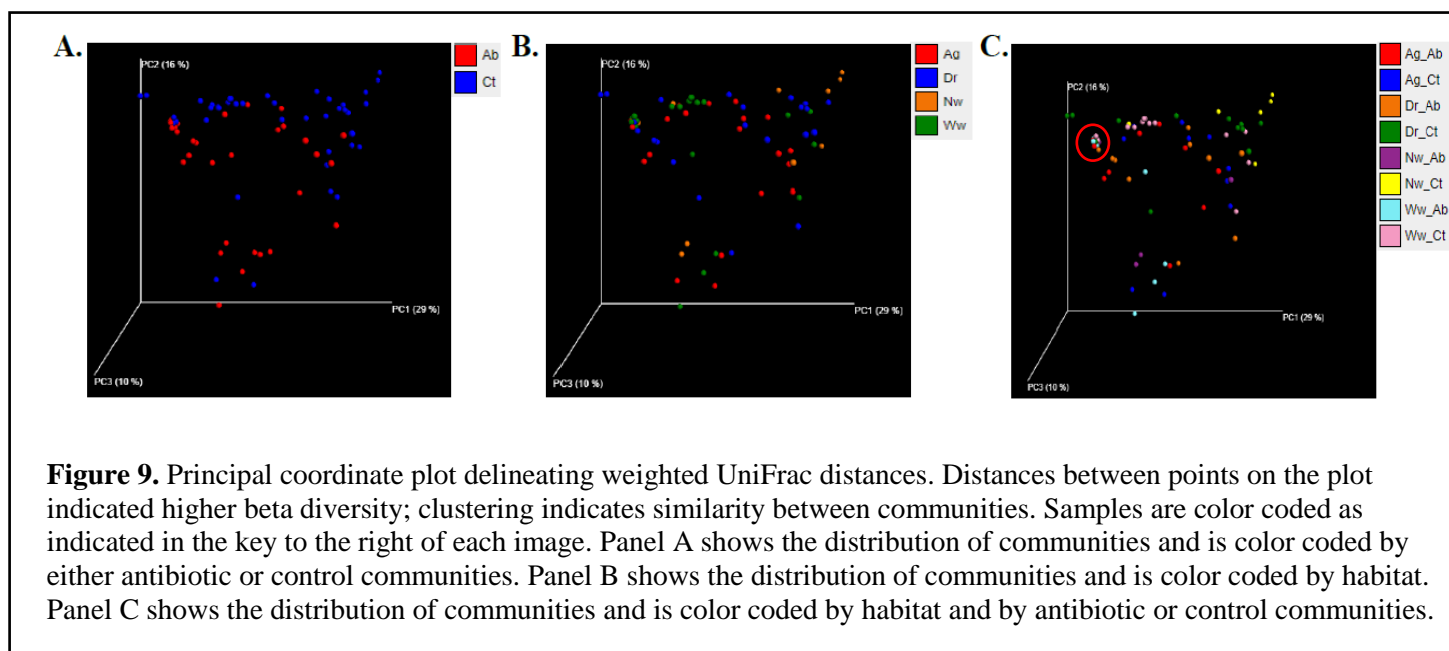
In summary, the alpha diversity analyses conducted revealed that the selection imposed by antibiotics on communities leads to decreased species diversity. The presence of antibiotics had

more profound impacts on diversity in agricultural waters and wastewaters, but presented negligible differences between natural waters and drain samples. Finally, agricultural waters were the most species rich of the habitats, which could also be due to nutrient richness due to the presence of manure, fertilizers, and even plants as nutrients.

Beta diversity, a measure of species dissimilarity between communities, was assessed using a weighted unique fraction metric (UniFrac) analysis in QIIME. In general, the UniFrac metric examines phylogenetic distances between various taxa and reflects the degree of similarity between differing communities — similar communities will cluster together in 3-dimensional space while dissimilar communities will be spaced further apart [46]. UniFrac metrics are categorized as unweighted (qualitative measures) or weighted (quantitative

measures). Weighted UniFrac was used in this study as it characterizes the distances, or dissimilarities, between communities as reflected by the numerical abundance of each taxa in that community as well as the phylogeny of the community members [47].

As in the alpha diversity analyses, samples were first compared for selective versus non-selective samples. The clustering of samples in principal coordinate analysis plots (PCAs) allows for a 3-dimensional rendering of the relative spatial distribution of each sample. For PCAs, clusters indicate high levels of similarity between samples, while spaced out samples are more dissimilar (and thus more diverse) from each other. In general, selective samples appeared closer together on the plot. Alternatively, non-selective samples showed a more variable distance range on the plot (**Figure 9A**).

Communities were then divided and analyzed based on habitat (**Figure 9B**). As an overall trend, agricultural and wastewater samples appeared closer to other members of their category, indicating lower beta diversity. Conversely, drain and natural water samples exhibited a greater range of coordinates on the plot, indicating higher beta diversity.

Figure 9. Principal coordinate plot delineating weighted UniFrac distances. Distances between points on the plot indicated higher beta diversity; clustering indicates similarity between communities. Samples are color coded as indicated in the key to the right of each image. Panel A shows the distribution of communities and is color coded by either antibiotic or control communities. Panel B shows the distribution of communities and is color coded by habitat. Panel C shows the distribution of communities and is color coded by habitat and by antibiotic or control communities.

Finally, sampling site categories were sub-divided into their selective and non-selective counterparts (**Figure 9C**). There was one large cluster containing 29 samples from all habitats, indicated by a red circle. This cluster indicates these communities were highly similar. Further analysis of the samples in this cluster revealed that all samples in the cluster contained *Ochrobactrum* sp. with 22 of the 29 clustered samples containing > 65% *Ochrobactrum* sp. Non-selective natural water samples appeared to have higher beta diversity, as they were the most variable in distance on the plot. Interestingly, non-selective drain samples showed similar diversity trends to the natural water non-selective samples. Most notably, wastewater antibiotic samples were very close together and thus very low in beta diversity.

Income and Diversity Comparisons

It was hypothesized that median household income would be correlated to diversity of antibiotic resistant bacteria; specifically, there would be a greater quantity of antibiotic resistant bacteria in areas of lower income as prior studies suggest [24,25]. Median household income was obtained for 47 different communities and further broken down into three main categories, high, medium, and low-income areas, as described in the Materials and Methods. The breakdown of income data into categories was as follows: 19 samples were collected from communities that were high income, 20 samples were collected from communities that were medium income, and 8 samples were collected from communities that were low income. Median household income for each community was plotted in comparison to the following diversity metrics: alpha diversity (Shannon diversity index), genera frequencies (both antibiotic and control groups), species frequencies (both antibiotic and control groups), pH, and temperature (**Figure 10**). All linear regressions resulted in p-values > 0.05 indicating that none of the pairings with income data were significant. Additionally, all R^2 values were very low (< 0.0), meaning the data was a poor fit to the regression line.

An ANCOVA analysis was used to further study the effect of the following independent co-variants: median household income, habitat, sample pH, and sample temperature on Shannon diversity index. As measured by squared-multiple R, the 4 co-variants explain 27.7% of variations in the Shannon diversity index between communities. However, of the 4 variables, only habitat and pH were significant ($p < 0.05$) in explaining differences in Shannon diversity indices. The p-value for habitat ($p = 0.036$) was smaller than the p-value for pH ($p = 0.050$), indicating that habitat has a stronger effect than pH on Shannon diversity.
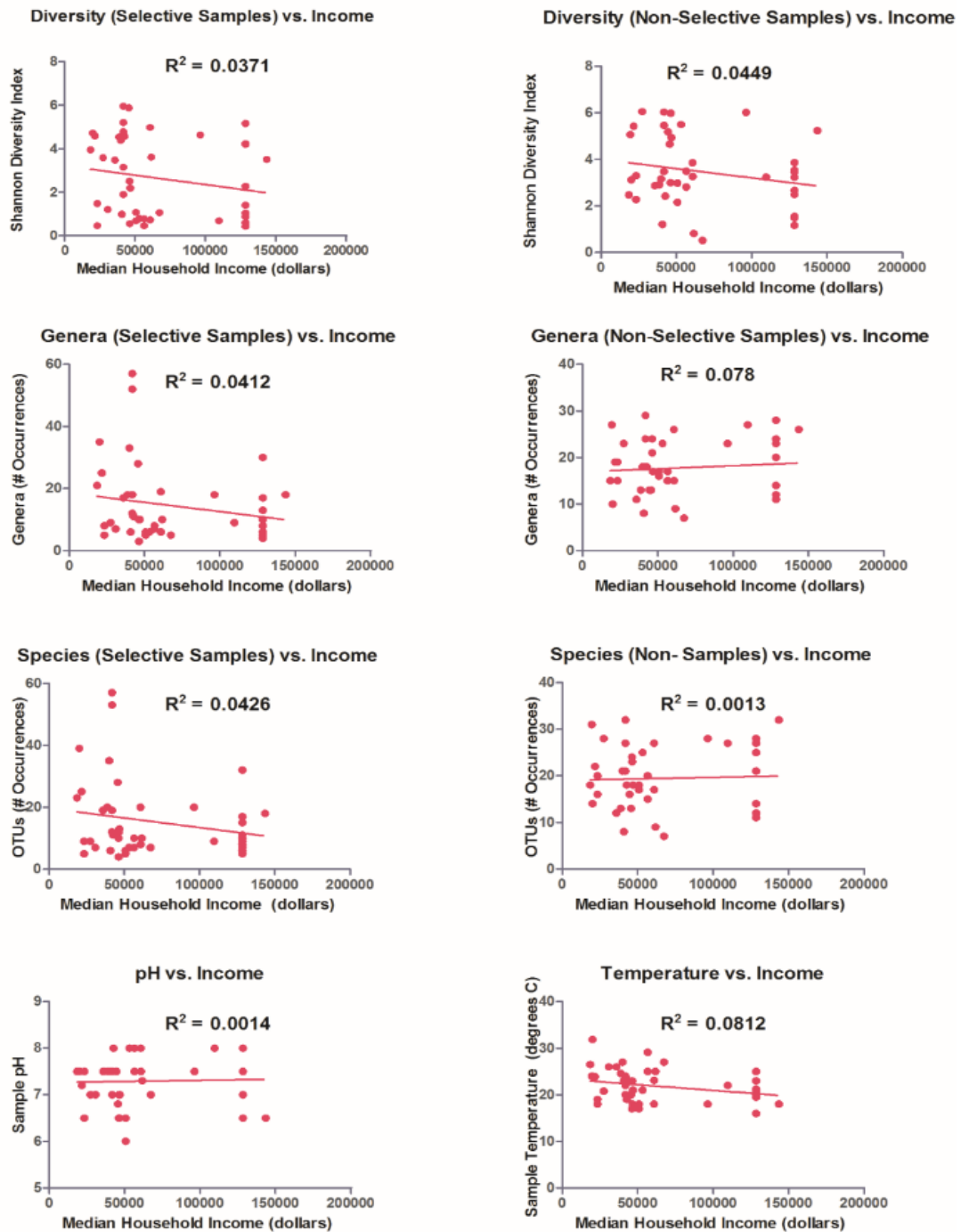
Figure 10. Median household income (in dollars) for each community plotted with the following metrics: Shannon diversity index, genus and species frequencies (selective and non-selective groups), temperature, and pH. All p-values > 0.05, indicating no significance between pairings.

Ochrobactrum *sp. Analysis*

The community metagenome analysis revealed the presence of *Ochrobactrum* sp. isolates in almost 90% of the communities (78 out of 87 communities). The vast predominance of *Ochrobactrum* sp. makes it clinically relevant to study this organism in more detail. *Ochrobactrum* is a genus of Gram-negative, rod-shaped, catalase positive bacteria, inhabiting primarily environmental soils and sediments, but has also been isolated from animal hosts, including humans [51–53]. Two of *Ochrobactrum'*s most studied species are *O. anthropi* and *O. intermedium* both of which are considered emerging opportunistic pathogens [43, 54–56]. With increasing reports of infections attributed to *Ochrobactrum* sp. isolates in immunocompromised individuals, it is imperative to obtain more knowledge on this genus. From a clinical standpoint, it is pertinent to determine to which antibiotics *Ochrobactrum* species are naturally resistant. Additionally, determining whether *Ochrobactrum* can spread its antibiotic resistance genes to a host bacterium is also important for better understanding the potential impacts of this human pathogen. To study these questions, *Ochrobactrum* isolates were obtained from samples dominated by this genus. We further identified a number of isolates from the genus *Pseudochrobactrum* to study as well. Both genera are part of the family, *Brucellaceae*, and 16S rRNA analysis clustered *Ochrobactrum* sp. and *Pseudochrobactrum* sp. closely together, but protein analysis revealed dissimilarities warranting distinct genera [57]. *Pseudochrobactrum* sp. have been reported as pathogenic to humans and have been isolated in humans, but reports of clinical isolation and are relatively rare [58,59]. Additionally, very few antibiotic resistant *Pseudochrobactrum* sp. have been isolated, except for one study describing sulfamethoxazole resistant *Pseudochrobactrum* sp. [60].

Phenotypic and 16S Sequencing Determination of *Ochrobactrum* Isolates

      To obtain *Ochrobactrum* sp. isolates, 10 communities were selected as the most likely candidates to contain the genus (based on the frequencies observed from metagenome sequencing results). Plating the communities on a medium that included vancomycin, colistin, and ceftazidime (to which *Ochrobactrum* species are known to be resistant) resulted in 18 single colony isolates. Differential tests were then conducted based on known *Ochrobactrum* phenotypic traits to narrow down the isolates further. Catalase testing did not eliminate any isolates as all isolates were catalase positive. However, Gram-staining removed six Gram-positive isolates from consideration. Genomic DNA preparations from the remaining 12 isolates were used as templates for PCR using universal 16S primers. The resultant purified PCR products were analyzed by 16S sequencing. These sequences were then compared to two databases to confirm their identity at the species level. Two samples could not be classified to the species level using either Genbank and RDP databases. Surprisingly, five of the isolates (from three different communities designated as FFRM2Ab, FFFPA2Ab, and FFRB2Ab) were classified as *Myroides* sp. In the metagenome study, *Myroides* sp. only appeared in eight communities. Additionally, in the metagenome study, though FFRB2Ab contained *Myroides* sp., neither FFFPA2Ab nor FFRM2Ab contained identified *Myroides* sp. isolates. 16S sequencing analyses revealed three of the twelve isolates were *Ochrobactrum* sp. Two of the *Ochrobactrum* sp. isolates were categorized to the species level. One isolate was classified as *Ochrobactrum anthropi* and one isolate was classified as *Ochrobactrum intermedium* (**Table 3**).

      In addition to the three *Ochrobactrum* sp. isolates that were obtained, three isolates were provided from Dr. Paul Himes's lab for further study (sample designations: S16, S19, and MR28). These isolates were also collected by the BIOL 501 students, but were sequenced and identified

by members of the Himes lab for 16S sequencing. Sequence analysis identified one isolate as *Ochrobactrum* sp. (isolate ID: S19) and two isolates (IDs: MR28 and S16) as *Pseudochrobactrum* sp. Thus, a collection of six related panel isolates were used to further analyses antibiotic resistance profiles.

Understanding *Ochrobactrum/Pseudochrobactrum* antibiotic resistance profiles

Upon obtaining *Ochrobactrum/Pseudochrobactrum* sp. isolates resistant to vancomycin and colistin (which exert their effects on cell wall and cell membrane synthesis, respectively), it was hypothesized that the isolates may be resistant to other antibiotics that also target cell wall or membrane synthesis, such as carbenicillin and imipenem. To determine levels of antibiotic resistance, minimum inhibitory concentration assays were performed on 3–6 replicates using seven different concentrations of eight different antibiotics. These assays allow for the identification the concentration at which a bacterial strain is resistant to the naked eye in a subjective manner. Because each strain grew at varying densities in the Mueller Hinton broth employed in this assay, the growth of each strain under selection was compared to control wells containing no antibiotics.

Sample Designation	Colony Appearance	Gram Stain	Catalase Test	Length (bp)	BLAST Designation	BLAST E-Value	RDP Classification	RDP Value
MFR1	white, opaque, circular	+	+	NA	NA	NA	NA	NA
MFR2	yellow, circular	+	+	NA	NA	NA	NA	NA
FFRM2-1	yellow, circular	-	+	759	*Myroides* sp.	0.0	*Myroides* sp.	100%
FFRM2-2-A	white, circular	+	+	NA	NA	NA	NA	NA

Table 3. Isolate Differentiation to Determine Presence of *Ochrobactrum* sp.

FFRM2-2-B	yellow, circular	+	+	NA	NA	NA	NA	NA
FFRM2-3	tan/white, opaque	-	+	679	*Myroides* sp.	0.0	*Myroides* sp.	100%
FFFPA2-1	yellow, circular	-	+	626	*Myroides oderatus*	0.0	*Myroides* sp.	100%
FFFPA2-2	small, white circular	+	+	NA	NA	NA	NA	NA
FFRB2-1-A	white, feathery	-	+	605	*Myroides* sp.	0.0	*Myroides* sp.	100%
FFRB2-1-B	yellow, circular	-	+	575	*Pseudomonas* sp.	2.00E-14	Unclassifed Bacteria	100%
FFRB2-2	yellow, circular	+	+	NA	NA	NA	NA	NA
FFRB2-3	clear	-	+	622	*Myroides* sp.	0.0	*Myroides* sp.	100%
GRS1-1	white, circular	-	+	672	*Rhizobium* sp.	7.00E-90	Alphaproteo bacteria	81%
CMG−	white, circular	-	+	718	*Ochrobactrum* sp.	0.0	*Ochrobactrum* sp.	95%
CMG+	white, circular	-	+	632	*Ochrobactrum intermedium*	0.0	*Ochrobactrum* sp.	71%
WHF1	yellow, small, circular	-	+	609	Uncultured bacterium	2.00E-54	*Unclassifed Bacteria*	88%
GRW1-1	white, opaque, circular	-	+	678	*Ochrobactrum anthropi*	0.0	*Ochrobactrum* sp.	89%
WMW1	white, circular	-	+	462	Could not be classified	NA	Could not be classified	NA

NA = Not applicable; indicated the isolate was not further tested after differential testing revealed the isolate was not Ochrobactrum

The following antibiotics were chosen for analysis: vancomycin, colistin, tetracycline, tobramycin, kanamycin, carbenicillin, trimethoprim, imipenem as they represent a variety of antibiotic classes. Low MICs indicate samples may be more sensitive to that particular antibiotic while high MICs indicate samples can tolerate higher concentrations of that particular antibiotic (**Table 4**). All samples showed more sensitivity to tetracycline and tobramycin than the other antibiotics. The former inhibited all samples at concentrations ≤ 1 µg/mL. The samples showed the most tolerance for vancomycin, colistin, imipenem, and carbenicillin, but there were a few exceptions: MR28 (*Pseudochrobactrum*) was more sensitive to carbenicillin, and S16

(*Pseudochrobactrum*) was more sensitive to colistin. Finally, intermediate tolerances (most samples had MIC values between 128–256 µg/mL) were observed for kanamycin and trimethoprim. One limitation of the traditional MIC assay is its sensitivity; up to 10^5 CFU/mL can be present in a sample but may still appear similar to sterile medium.

Table 4. MIC values for *Ochrobactrum*/*Pseudochrobactrum* isolates									
		2MIC values (µg/mL) [†]							
Genus/Species	Isolate Designation	Vancomycin	Colistin	Tetracycline	Imipenem	Kanamycin	Carbenicillin	Tobramycin	Trimethoprim
Ochrobactrum species									
Ochrobactrum anthropi	GRW1-1	512	512	0.5	512	≥128	512	2	≥128
Ochrobactrum sp.	CMG−	512	512	1	512	≥256	512	≥8	≥128
Ochrobactrum intermedium	CMG+	512	512	1	512	≥256	512	≥8	≥128
Ochrobactrum sp.	S19	≥256	512	1	512	≥128	512	4	512
Pseudochrobactrum species									
Pseudochro-bactrum sp.	MR28	≥256	≥256	1	≥128	≥32	≥8	8	≥32
Pseudochro-bactrum sp.	S16	≥128	≥4	0.5	≥256	≥32	≥256	0.5	≥128
[†]two trials of triplicate assays for antibiotic concentrations from 0–512 µg/mL and one trial of triplicate assays for antibiotic concentrations from 0–32 µg/mL. "≥" symbol indicates MICs values that may be at or above this level due to replicates that showed differing MICs.									

Strains were further analyzed using a more quantitative method, spectrophotometry, to examine the growth of isolates in response to differing doses of each antibiotic. Normalizing the data according to growth in non-selective conditions allowed for the comparison of quantitative MIC values across all isolates. For this, each strain's O.D.$_{595}$ reading was converted to a percentage of its non-selective growth and was graphed as a function of antibiotic concentration (Note: there were two different trials of antibiotic concentrations conducted. One trial analyzed higher concentrations of antibiotics from 0–512 µg/mL, while another trial analyzed lower concentrations

of antibiotics for 0–32 μg/mL) (**Figures 11 and 12**). In general, a downward sloping line was observed for samples sensitive to increased antibiotic concentration. This observation was apparent for tetracycline and tobramycin. A more linear correlation was observed when samples were less affected by increasing antibiotic concentration as was observed for vancomycin, colistin, imipenem, carbenicillin, and trimethoprim. Kanamycin showed a different trend in which growth was unaffected at low antibiotic concentrations and gradually decreased at the highest antibiotic concentration.

A few conclusions can be drawn from the quantitative MIC data. Two samples, CMG– (*Ochrobactrum* sp.) and CMG+ (*O. intermedium*), consistently behaved similarly across all trials. These samples were from the same geographic community, but one was obtained from a non-selective sample (*Ochrobactrum* sp. CMG–), and the other from a selective sample (*O. intermedium* CMG+). For a few of the antibiotics examined, MR28 (*Pseudochrobactrum* sp) and GRW1-1 (*O. anthropi*) often showed more growth than other samples, but the trend was inconsistent across multiple replicates. Of note, some samples' large standard deviations indicated a high variability between replicates that may reflect technical errors or true heterogeneity. Another interesting observation, MR28 and S16, the two *Pseudochrobactrum* sp. isolates, differed in their growth even under non-selective conditions. S16 was consistently slow growing compared to MR28. However, in general, both *Pseudochrobactrum* sp. isolates typically were slightly more sensitive to antibiotics than the *Ochrobactrum* sp. isolates. Finally, trends between the mode of antibiotic action and the whole panel's relative sensitivities were observed. For example, tetracycline and tobramycin target the translation step of protein synthesis and were particularly deleterious to the entire panel's growth. The samples were very tolerant to high concentrations of vancomycin, colistin, imipenem and carbenicillin which all
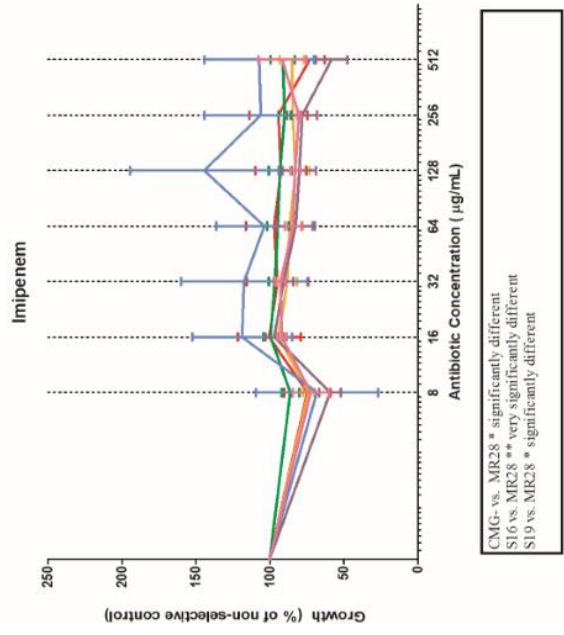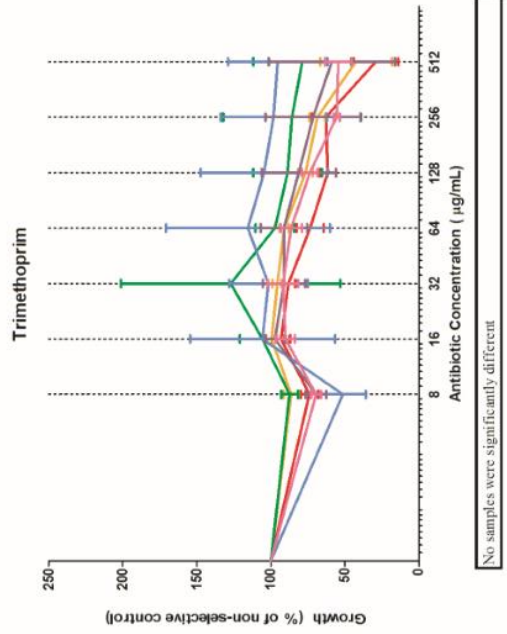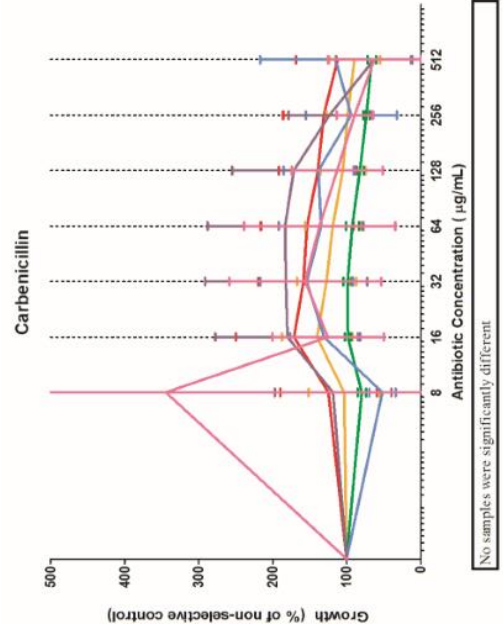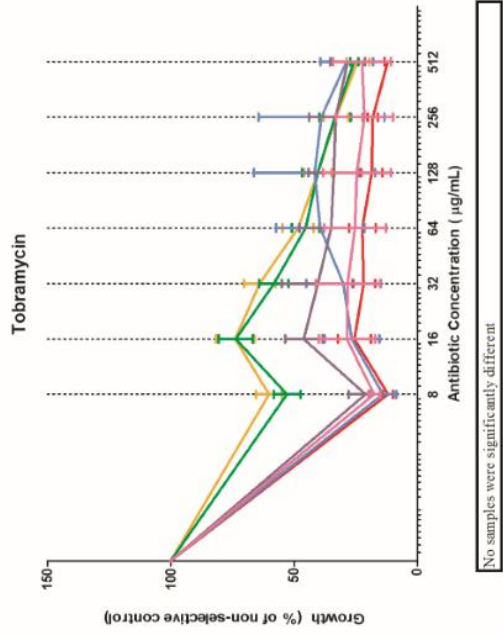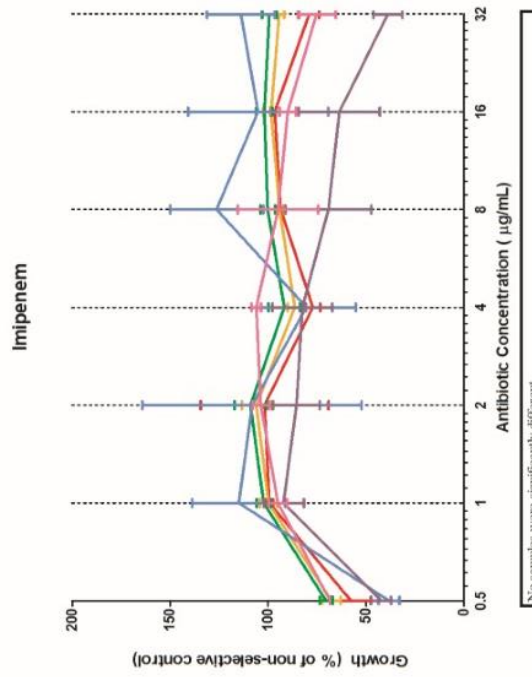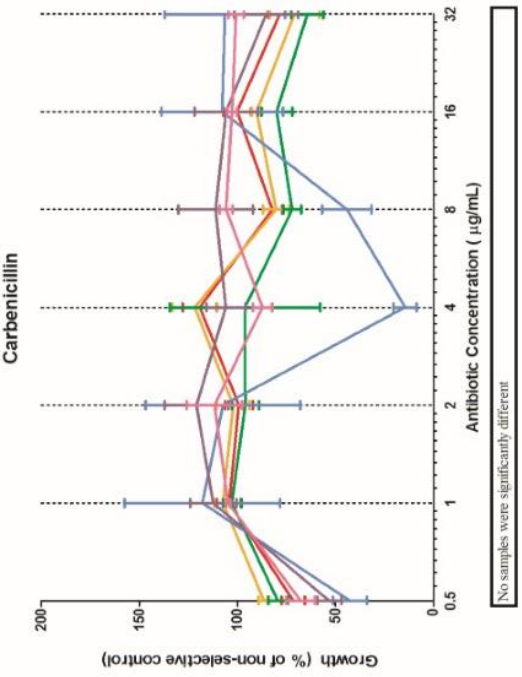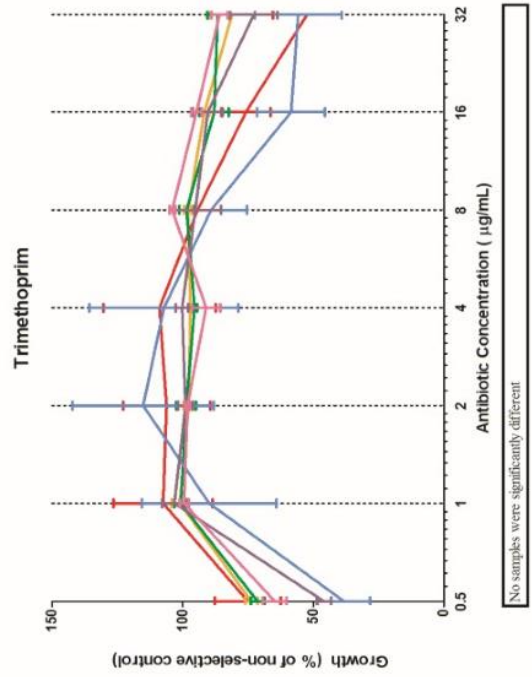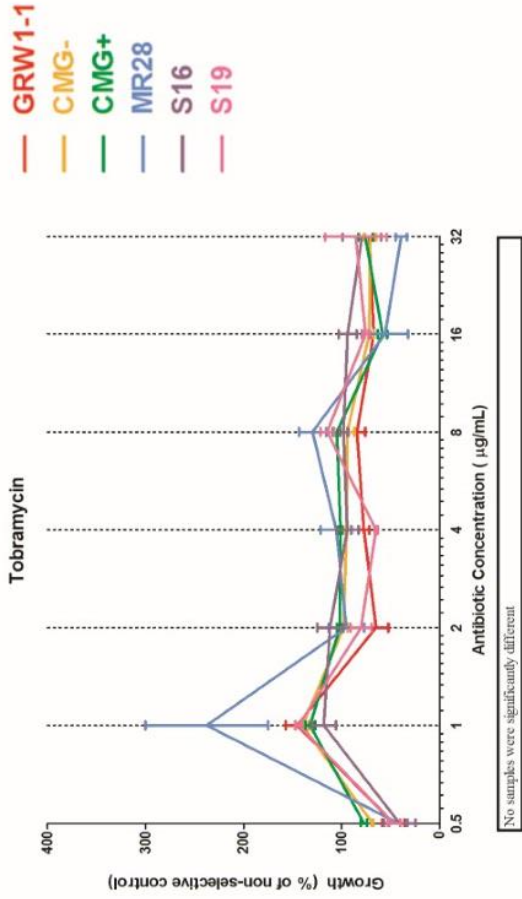
Figure 11. MICs for concentrations between 0 - 512 µg/mLas measured by spectrophotometry, O.D.595. Growth at each antibiotic concentration was normalized as a percentage of the growth that was observed for each isolate with no antibiotics. Significances were determined using one-way ANOVA analysis with Tukey's multiple comparisons post-test. Asterisks indicate the following significances: *** - p < 0.001; ** - 0.001 < p < 0.01; * - 0.01 < p < 0.05

52
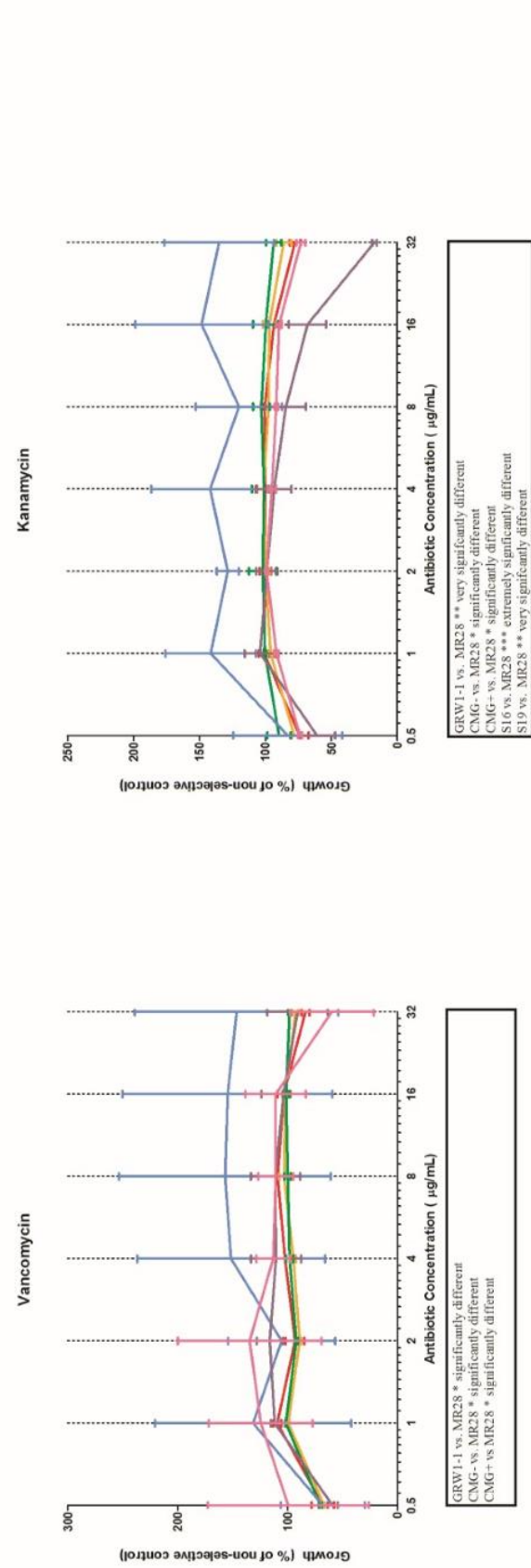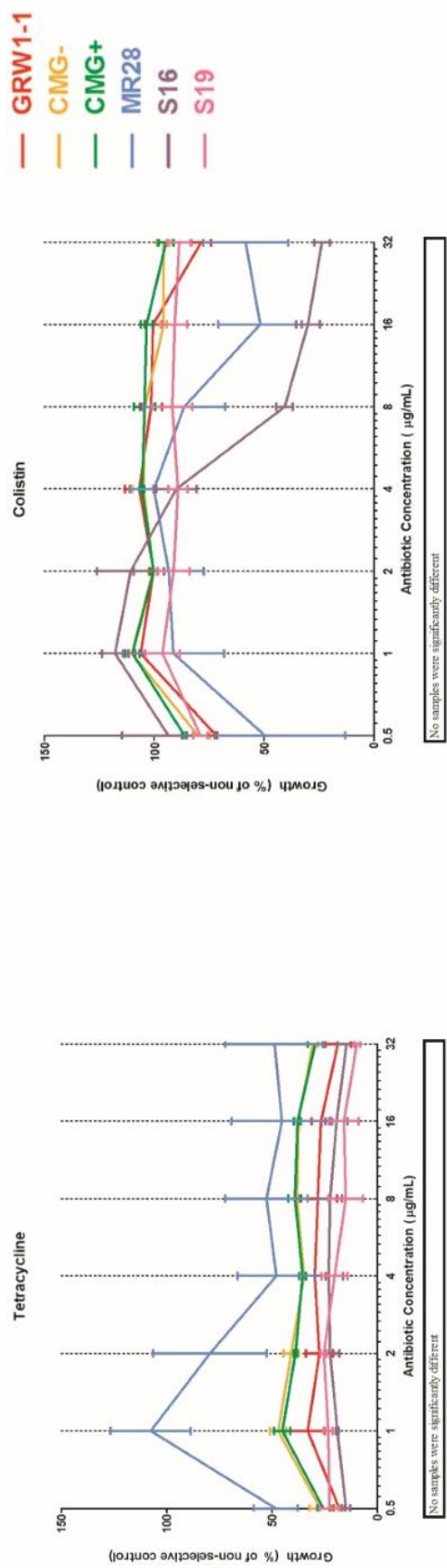
Figure 12. MICs for concentrations between 0 - 32 μg/mL as measured by spectrophotometry, O.D.595. Growth at each antibiotic concentration was normalized as a percentage of the growth that was observed for each isolate with no antibiotics. Significances were determined using one-way ANOVA analysis with Tukey's multiple comparisons post-test. Asterisks indicate the following significances: *** - p < 0.001; ** - 0.001 < p < 0.01; * - 0.01 < p < 0.05

Figure 11. MICs for concentrations between 0 - 32 µg/mL as measured by spectrophotometry, O.D.595. Growth at each antibiotic concentration was normalized as a percentage of the growth that was observed for each isolate with no antibiotics. Significances were determined using one-way ANOVA analysis with Tukey's multiple comparisons post-test. Asterisks indicate the following significances: *** - $p < 0.001$; ** - $0.001 < p < 0.01$; * - $0.01 < p < 0.05$

inhibit portions of cell wall synthesis, leading to speculation that the panel may have a change in their cell walls that diminished the effects of cell wall targeting antibiotics.

Transmission of antibiotic resistance by Ochrobactrum

Horizontal gene transfer between bacterial cells plays a major role in the spread of genes conferring antibiotic resistance. It was hypothesized that *Ochrobactrum* may be able to transfer colistin resistance to another bacterial species by conjugation. Using a kit-based plasmid purification procedure, the extraction of pure plasmid was attempted from the six panel *Ochrobactrum* and *Pseudochrobactrum* isolates. However, spectrophotometry indicated low yields of total DNA (<10 ng/µL). This result indicates the following possibilities: the isolates did not contain a plasmid, the isolates maintain their plasmids at a low copy number, or the isolates did not lyse during the extraction procedure. Since two of these explanations posit that the isolates contained plasmids, it was pertinent to determine whether colistin resistance could be passed from a chosen panel isolate to a colistin-sensitive bacterial strain of *Pseudomonas aeruginosa*.

The *Ochrobactrum* sp. isolate CMG– (which was consistently tolerant of high doses of colistin) was spotted with a colistin-sensitive, recipient strain, *P. aeruginosa* PAO1, at high densities using a routine method for conjugation of *P. aeruginosa* with *E.coli* donor strains [61– 63]. It was noted that plating the bacteria specimens together resulted in a color change in the densely populated mixed species spots. On LB agar plates, *Ochrobactrum* sp. CMG– appears as a creamy, white color, while *P. aeruginosa* PAO1 is light yellow in color. However, the mixed spots were bright greenish–blue in color, as can be seen in **Supplementary Figure 1** in the Appendix [64]. This color change is likely due to the production of two pigments, pyocyanin and pyoverdine, that function in competition and iron acquisition by *P. aeruginosa*. Interestingly, upon plating the mating mixture of *Ochrobactrum* sp. CMG– and *P. aeruginosa* on selective plates, no colonies

grew. With no counter-selection applied, the colistin-resistant *Ochrobactrum* sp. present should have still grown, but must have died during this process. Because *P. aeruginosa* also did not grow across several trials, it can be concluded that this *Ochrobactrum* sp. isolate was unable to transfer its resistance to *P. aeruginosa*.

This experiment was repeated to test for the transfer of vancomycin resistance from *Ochrobactrum* sp. CMG– to *P. aeruginosa* PAO1. However, *P. aeruginosa* was inherently resistant to vancomycin and it could not be determined whether *Ochrobactrum* sp. CMG– could transfer vancomycin resistance to a host bacterium. Future efforts could include using a counter-selective method to eliminate *Ochrobactrum* sp. after mating and testing additional isolates or utilizing other vancomycin-sensitive recipient strains.

Discussion

The initial goal of this study was to isolate and identify antibiotic resistant bacteria in Louisville, KY. Sequencing of whole community metagenomes revealed the presence of 116 different bacterial genera, 86 of which came from selective communities. This result may be clinically significant because it could indicate that a multitude of bacteria are capable of manifesting antibiotic resistance, either inherently or by acquiring resistance mechanisms from other bacteria. A wide array of antibiotic resistant bacteria is of concern because they may cause a variety of infections that are difficult to treat.

Previous studies indicate the diversity of bacteria found in water environments and indicated the conduciveness of water-based ecosystems to the spread of antibiotic resistance [7,10]. However, there was little knowledge of the microbial ecology in local water environments. Thus, this study also aimed to explore the bacterial diversities of four different local environmental

water sources (natural waterways, agricultural waters, commercial drain waters, and wastewaters). Agricultural waters and commercial drain areas showed the greatest levels of species richness, indicating the possibility that bacterial diversity increases in areas with more human interference. For example, humans may alter agricultural areas by modifying land for crop production, spraying pesticides, and administering antimicrobial agents, and humans alter commercial drain habitats with the use of cleaning agents. Additionally, natural waters (which likely are areas of lower human interference) showed the lowest levels of species richness. The lower species richness in wastewater samples was surprising, but perhaps the composition of this environment (such as availability of nutrients) greatly favored a select few bacteria, inducing competition and lowering species richness.

Additionally, this study aimed to analyze the possible relationship between community affluence and antibiotic resistant bacteria. It was hypothesized that lower income areas would have more diversity of antibiotic resistant bacteria, due to financial strains leading to improper antibiotic use [24,25]. However, linear regressions and ANCOVA analysis revealed that income was not a significant predictor of diversity. However, we speculate that this could be due to a potential sampling bias in terms of total number of samples, the uneven distribution of high, medium, and low-income sites, or geographic biases. For example, many communities were clustered in the central part of Louisville and the eastern outskirts of the city, but no samples were taken from the far west side of the city. Additionally, there was not an even spread of income data. Frequencies of high and medium categories were nearly equivalent, but the low category was significantly less frequent than the other categories. For future studies, sampling should be more evenly spread; for example, by aiming to obtain equal quantities of samples from every zip-code in Louisville to ensure each area of the city is represented.

In addition to a possible sampling bias, there are a few other factors that may have influenced the results of this study. For instance, isolating bacteria by culture-dependent methods may have limited the growth of some bacteria which would reduce overall diversity [65– 67]. The primary antibiotics studied (vancomycin and colistin) also may have impacted results. For example, vancomycin targets peptidoglycan synthesis, which would be more detrimental to Gram-positive organisms; thus, some Gram-negative organisms that grew under selective conditions, may have only been impacted by colistin. Conversely, Gram-positive organisms would have been primarily affected by vancomycin, as colistin targets the outer membrane, found only in Gram-negative bacteria. Finally, there was a notable discrepancy between 16S individual isolate sequencing and community metagenome sequencing, in that *Ochrobactrum* sp. dominated at the community level (appearing in 78 of 87 communities), but only one individual isolate was identified as *Ochrobactrum* sp. in the original data set. Though many isolates originally were not identified by 16S sequencing as *Ochrobactrum* species, members of Dr. Paul Himes's lab have since re-prepared and re-sequenced those isolates and only identified one additional isolate as *Ochrobactrum* sp. With a somewhat non-descript physical appearance (white, circular colonies) it is possible many *Ochrobactrum* sp. isolates blended in with other colonies and were missed upon initial single colony isolation or a sampling bias occurred where more colorful colonies were chosen over white ones.

With the prevalence of *Ochrobactrum* sp. in both selective and non-selective communities and the genera's status as an emerging opportunistic pathogen, it is important to learn more about the organism [68]. One species in this genus, *Ochrobactrum anthropi* is classified in the literature as an opportunistic pathogen and causes infection in immunocompromised and patients with catheters; however, the species has also infected seemingly healthy individuals [54,55]. It was

58

further shown that another species in this genus, *Ochrobactrum intermedium*, may bbe commonly misidentified as *O. anthropi*, causing speculation as to whether other infections initially attributed to *O. anthropi* may have instead been caused by *O. intermedium* [43]. Supporting that speculation and its status as an emerging opportunistic pathogen, *O. intermedium* has had numerous reports of its clinical isolation, including: in a patient with bladder cancer which resulted in bacteremia; in a hemodialysis patient which resulted in infective endocarditis; and in a liver transplant patient which resulted in abscesses [56,69,70]. In addition to immunocompromised human hosts, *Ochrobactrum* sp. isolates have also been found in animal hosts. These organisms have been collected in the fecal matter from chickens and turkeys, and both hosts were from commercial agriculture sites [52,53]. Isolating bacteria from animal fecal matter has implications in that these strains are exposed to the environment and may be easily transferrable to other habitats by other animal carriers or by run-off into water sources.

O. intermedium's habitat is described in literature as a "human-associated technological niche" as it has been found primarily in areas with large amounts of human activity and interference in the natural environment (especially associated with polluted areas) [71]. As such, urban areas, like Louisville, are likely to be prime breeding grounds for *O. intermedium*. Additionally, agricultural areas outside of the city may also be *O. intermedium* hotspots and due to the previously discussed high antibiotic use, may also be prominent areas for *O. intermedium* to develop resistance. Interestingly, there have been relatively few reports of *O. intermedium* isolates from water environments [71]. Therefore, results from this study with several water isolates of potentially *O. intermedium* could be particularly noteworthy.

Antibiotic resistance trends for *Ochrobactrum* sp. have been studied previously. Most studies agree that all *Ochrobactrum* sp. are inherently resistant to β-lactam and carbapenem

antibiotics, with the exception of imipenem [44,72]. Another study concluded widespread susceptibility in the genus to trimethoprim [44,54]. These observations notably contrast with results from this study, wherein all panel isolates were uninhibited by imipenem at concentrations < 128 μg/mL. Additionally, all *Ochrobactrum* sp. isolates were uninhibited by trimethoprim at concentrations < 128 μg/mL (one *Pseudochrobactrum* sp. isolate was inhibited by the antibiotic at low concentrations, 32 μg/mL). Differences in resistance amongst different *Ochrobactrum* species has been reported; mainly, *O. anthropi* was reported to be susceptible to colistin and tobramycin while *O. intermedium* was resistant to these antibiotics [44]. Though some studies have been conducted on these organisms, confounding results have been reported. Additional research is required to obtain sure mechanisms to differentiate *Ochrobactrum* sp. isolates [44].

Overall, from this study, it can be concluded that water environments are significant sources of antibiotic resistant bacteria, and aquatic habitats located in the vicinity of commercial agriculture areas may be especially species rich. *Ochrobactrum* sp. appeared frequently in the communities sampled in this study, perhaps indicating this bacterial genus thrives in these areas and may utilize aquatic habitats to acquire antibiotic resistance. Based on the results of this project, future directions could include the expansion at sampling sites to achieve an even distribution to achieve a better reflection of bacterial ecology across Louisville. Continued study of *Ochrobactrum*, such as testing its growth against additional antibiotics and testing its ability to transfer other antibiotic resistances to additional sensitive recipients, is necessary to develop a broader understanding of the clinical and public health significance of antibiotic resistant *Ochrobactrum* species.

Acknowledgements

References

1.   Ventola, C. L. (2015). "The Antibiotic Resistance Crisis: Part 1: Causes and Threats." <u>Pharmacy and Therapeutics</u> **40**(4): 277-283.

2.   Chambers, H. F. (2001). "The changing epidemiology of *Staphylococcus aureus*?" <u>Emerg Infect Dis</u> **7**(2): 178-182.

3.   Spellberg, B. and D. N. Gilbert (2014). "The Future of Antibiotics and Resistance: A Tribute to a Career of Leadership by John Bartlett." <u>Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America</u> **59**(Suppl 2): S71-S75.

4.   Klevens, R., et al. (2007). "Invasive methicillin-resistant *Staphylococcus aureus* infections in the united states." <u>JAMA</u> **298**(15): 1763-1771.

5.   Koch, G., et al. (2014). "Evolution of Resistance to a Last-Resort Antibiotic in *Staphylococcus aureus* via Bacterial Competition." <u>Cell</u> **158**(5): 1060-1071.

6.   McKenna, M. (2013). "Antibiotic resistance: The last resort." <u>Nature</u> **499**(7459): 394-396.

7.   Llirós, M., et al. (2014). "Bacterial Community Composition in Three Freshwater Reservoirs of Different Alkalinity and Trophic Status." <u>PLoS ONE</u> **9**(12): 1-27.

8.   Sugumar, R. and B. Anandharaj (2016). "Assessment of Bacterial Load in the Fresh Water Lake System of Tamil Nadu." <u>International Journal of Current Microbiology and Applied Sciences</u> **5**(6): 236-246.

9.   Xin, W., et al. (2007). "Bacterial community composition of a shallow hypertrophic freshwater lake in China, revealed by 16S rRNA gene sequences." <u>FEMS Microbiology Ecology</u> **61**(1): 85-96.

10.   Marti, E., et al. "The role of aquatic ecosystems as reservoirs of antibiotic resistance." <u>Trends in Microbiology</u> **22**(1): 36-41.

11.   Davies, J. E. (1997). "Origins, acquisition and dissemination of antibiotic resistance determinants." <u>Ciba Found Symp</u> **207**: 15-27; discussion 27-35.

12.   Bartlett, J. G., et al. (2013). "Seven Ways to Preserve the Miracle of Antibiotics." <u>Clinical Infectious Diseases</u> **56**(10): 1445-1450.

13.   Marshall, B. M. and S. B. Levy (2011). "Food animals and antimicrobials: impacts on human health." <u>Clin Microbiol Rev</u> **24**(4): 718-733.

14. Fleming-Dutra, K. E., et al. (2016). "Prevalence of Inappropriate Antibiotic Prescriptions Among US Ambulatory Care Visits, 2010-2011." <u>JAMA: Journal of the American Medical Association</u> **315**(17): 1864-1873.

15. Davies, J., et al. (2006). "The world of subinhibitory antibiotic concentrations." <u>Current Opinion in Microbiology</u> **9**(5): 445-453.

16. Guerin, É., et al. (2009). "The SOS Response Controls Integron Recombination." <u>Science</u> **324**(5930): 1034-1034.

17. Kumar, K., C. Gupta, S., Chander, Y., & Singh, A. K. (2005). "Antibiotic Use in Agriculture and Its Impact on the Terrestrial Environment." <u>Advances in Agronomy</u> **87**: 1-54.

18. Watanakunakorn, C. (1984). "Mode of action and in-vitro activity of vancomycin." <u>J Antimicrob Chemother</u> **14 Suppl D**: 7-18.

19. Cetinkaya, Y., et al. (2000). "Vancomycin-resistant enterococci." <u>Clin Microbiol Rev</u> **13**(4): 686-707.

20. Kuzin, A. P., et al. (2000). "Enzymes of vancomycin resistance: the structure of D-alanine-D-lactate ligase of naturally resistant *Leuconostoc mesenteroides*." <u>Structure</u> **8**(5): 463-470.

21. Loho, T. and A. Dharmayanti (2015). "Colistin: an antibiotic and its role in multiresistant Gram-negative infections." <u>Acta Med Indones</u> **47**(2): 157-168.

22. Yahav, D., et al. (2012). "Colistin: new lessons on an old antibiotic." <u>Clin Microbiol Infect</u> **18**(1): 18-29.

23. Liu, Y.-Y. (2016). "Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study." <u>Lancet Infectious Diseases, The</u> **16**(2): 161-168.

24. McKee, M. D., et al. (1999). "Antibiotic Use for the Treatment of Upper Respiratory Infections in a Diverse Community." <u>Journal of Family Practice</u> **48**(12): 993-996.

25. Richman, P. B., et al. "Oral antibiotic use without consulting a physician: A survey of ED patients." <u>The American Journal of Emergency Medicine</u> **19**(1): 57-60.

26. Planta, M. B. (2007). "The role of poverty in antimicrobial resistance." <u>J Am Board Fam Med</u> **20**(6): 533-539.

27. Pehrsson, E. C., et al. (2016). "Interconnected microbiomes and resistomes in low-income human habitats." Nature **533**: 212.

28. Vingataramin, L. and E. H. Frost (2015). "A single protocol for extraction of gDNA from bacteria and yeast." Biotechniques **58**(3): 120-125.

29. Desjardins, P. and D. Conklin (2010). "NanoDrop Microvolume Quantitation of Nucleic Acids." Journal of Visualized Experiment: JoVE(45): 2565.

30. Clarridge, J. E. (2004). "Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases." Clin Microbiol Rev **17**(4): 840-862.

31. Böttger, E. C. (1989). "Rapid determination of bacterial ribosomal RNA sequences by direct sequencing of enzymatically amplified DNA." FEMS Microbiology Letters **65**(1-2): 171-176.

32. Woese, C. R., et al. (1985). "A Phylogenetic Definition of the Major Eubacterial Taxa." Systematic and Applied Microbiology **6**(2): 143-151.

33. Hongoh, Y., et al. (2003). "Evaluation of primers and PCR conditions for the analysis of 16S rRNA genes from a natural environment." FEMS Microbiol Lett **221**(2): 299-304.

34. Hall, T. A. (1999). "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." Nucleic Acids Symposium Series **41**: 95-98.

35. Altschul, S. F., et al. (1990). "Basic local alignment search tool." Journal of Molecular Biology **215**(3): 403-410.

36. Cole, J. R., et al. (2014). "Ribosomal Database Project: data and tools for high throughput rRNA analysis." Nucleic Acids Research **42**(D1): D633-D642.

37. Wilson, Micheal C. and J. Piel (2013). "Metagenomic Approaches for Exploiting Uncultivated Bacteria as a Resource for Novel Biosynthetic Enzymology." Chemistry & Biology **20**(5): 636-647.

38. Caporaso, J. G., et al. (2010). "QIIME allows analysis of high-throughput community sequencing data." Nature Methods **7**: 335.

39. Kumar, R., et al. (2014). "Getting Started with Microbiome Analysis: Sample Acquisition to Bioinformatics." Current protocols in human genetics / editorial board, Jonathan L. Haines ... [et al.] **82**: 18.18.11-18.18.29.

40. Schloss, P. D. and S. L. Westcott (2011). "Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis." Appl Environ Microbiol **77**(10): 3219-3226.

41. Vazquez-Baeza, Y., et al. (2013). "EMPeror: a tool for visualizing high-throughput microbial community data." Gigascience **2**(1): 16

42. Higgins, C. S., et al. (2001). "Resistance to antibiotics and biocides among non-fermenting Gram-negative bacteria." Clinical Microbiology & Infection **7**(6): 308-315.

43. Velasco, J., et al. (1998). "Evaluation of the relatedness of *Brucella* spp. and *Ochrobactrum anthropi* and description of *Ochrobactrum intermedium* sp. nov., a new species with a closer relationship to *Brucella* spp." Int J Syst Bacteriol **48 Pt 3**: 759-768.

44. Teyssier, C., et al. (2005). "Molecular and phenotypic features for identification of the opportunistic pathogens *Ochrobactrum* spp." J Med Microbiol **54**(Pt 10): 945-953.

45. Whittaker, R. H. (1972). "Evolution and Measurement of Species Diversity." Taxon **21**(2/3): 213-251.

46. Lozupone, C. and R. Knight (2005). "UniFrac: a new phylogenetic method for comparing microbial communities." Appl Environ Microbiol **71**(12): 8228-8235.

47. Lozupone, C. A., et al. (2007). "Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities." Appl Environ Microbiol **73**(5): 1576-1585.

48. Fierer, N., et al. (2012). "Cross-biome metagenomic analyses of soil microbial communities and their functional attributes." Proceedings of the National Academy of Sciences **109**(52): 21390.

49. Castañeda, L. E. and O. Barbosa (2017). "Metagenomic analysis exploring taxonomic and functional diversity of soil microbial communities in Chilean vineyards and surrounding native forests." PeerJ **5**: e3098.

50. Colwell, R. K., et al. (2012). "Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages." Journal of Plant Ecology **5**(1): 3-21.

51. Lebuhn, M., et al. (2000). "Taxonomic characterization of *Ochrobactrum* sp. isolates from soil samples and wheat roots, and description of *Ochrobactrum tritici* sp. nov. and *Ochrobactrum grignonense* sp. nov." Int J Syst Evol Microbiol **50 Pt 6**: 2207-2223.

52. ElAdawy, H., et al. (2012). "Isolation and characterization of *Ochrobactrum anthropi* and *Ochrobactrum pecoris* from caecal content of commercial turkeys." Veterinary Microbiology **155**(2-4): 349-354.

53. Kämpfer, P., et al. (2003). "Towards a standardized format for the description of a novel species (of an established genus): *Ochrobactrum gallinifaecis* sp. nov." International Journal of Systematic and Evolutionary Microbiology **53**(3): 893-896.

54. Thoma, B., et al. (2009). "Identification and antimicrobial susceptibilities of *Ochrobactrum* spp." International Journal of Medical Microbiology **299**(3): 209-220.

55. Hagiya, H., et al. (2013). "Clinical characteristics of *Ochrobactrum anthropi* bacteremia." J Clin Microbiol **51**(4): 1330-1333.

56. Apisarnthanarak, A., et al. (2005). "Evaluation *of Ochrobactrum intermedium* bacteremia in a patient with bladder cancer." Diagnostic Microbiology and Infectious Disease **53**(2): 153-155.

57. Scholz, H. C., et al. (2008). "Genetic diversity and phylogenetic relationships of bacteria belonging to the *Ochrobactrum–Brucella* group by *recA* and 16S rRNA gene-based comparative sequence analysis." Systematic and Applied Microbiology **31**(1): 1-16.

58. Kämpfer, P., et al. (2006). "Description of *Pseudochrobactrum* gen. nov., with the two species *Pseudochrobactrum asaccharolyticum* sp. nov. and *Pseudochrobactrum saccharolyticum* sp. nov." International Journal of Systematic and Evolutionary Microbiology **56**(8): 1823-1829.

59. Seng, P., et al. (2013). "Identification of Rare Pathogenic Bacteria in a Clinical Microbiology Laboratory: Impact of Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry." Journal of Clinical Microbiology **51**(7): 2182-2194.

60. Gao, P., et al. (2012). "Occurrence of sulfonamide and tetracycline-resistant bacteria and resistance genes in aquaculture environment." Water Research **46**(7): 2355-2364.

61. Carter, M. Q., et al. (2010). "The *Pseudomonas aeruginosa* pathogenicity island PAPI-1 is transferred via a novel type IV pilus." J Bacteriol **192**(13): 3249-3258.

62. Kulasekara, B. R., et al. (2006). "Acquisition and evolution of the *exoU* locus in *Pseudomonas aeruginosa*." J Bacteriol **188**(11): 4037-4050.

63.   Hoang, T. T., et al. (2000). "Integration-Proficient Plasmids for *Pseudomonas aeruginosa*: Site-Specific Integration and Use for Engineering of Reporter and Expression Strains." Plasmid **43**(1): 59-72.

64.   Cox, C. D. and P. Adams (1985). "Siderophore activity of pyoverdin for *Pseudomonas aeruginosa*." Infection and Immunity **48**(1): 130-138.

65.   Vaz-Moreira, I., et al. (2011). "Culture-dependent and culture-independent diversity surveys target different bacteria: a case study in a freshwater sample." Antonie Van Leeuwenhoek **100**(2): 245-257.

66.   Stefani, F. O. P., et al. (2015). "Culture-Dependent and -Independent Methods Capture Different Microbial Community Fractions in Hydrocarbon-Contaminated Soils." PLoS ONE **10**(6): e0128272.

67.   Dickson, R. P., et al. (2014). "Analysis of Culture-Dependent versus Culture-Independent Techniques for Identification of Bacteria in Clinically Obtained Bronchoalveolar Lavage Fluid." Journal of Clinical Microbiology **52**(10): 3605-3613.

68.   Edmond, M. B., et al. (1999). "Nosocomial Bloodstream Infections in United States Hospitals: A Three-Year Analysis." Clinical Infectious Diseases **29**(2): 239.

69.   Moller, L. V., et al. (1999). "*Ochrobactrum intermedium* infection after liver transplantation." J Clin Microbiol **37**(1): 241-244.

70.   Bharucha, T., et al. (2017). "*Ochromobactrum intermedium*: an emerging opportunistic pathogen-case of recurrent bacteraemia associated with infective endocarditis in a haemodialysis patient." New Microbes New Infect **15**: 14-15.

71.   Aujoulat, F., et al. (2014). "Niches, Population Structure and Genome Reduction in *Ochrobactrum intermedium*: Clues to Technology-Driven Emergence of Pathogens." PLoS ONE **9**(1): 1-14.

72.   Alonso, C. A., et al. (2017). "Diversity of *Ochrobactrum* species in food animals, antibiotic resistance phenotypes and polymorphisms in the *blaOCH* gene." FEMS Microbiol Lett **364**(17).

73.   Centers for Disease Control and Prevention. (2013). *Antibiotic Resistance Threats in the Unites States*. Retrieved from: https://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf

Appendix

I. Supplementary Protocol 1

II. Supplementary Figure 1

I. Supplementary Protocol

Protocol for analyzing paired-end reads of 16S metagenomes using QWRAP (Prepared by Dr. Deborah Yoder-Himes)

Preparing the data and programs

Programs: Download and install QIIME v. 1.8, QWRAP v. 2, USEARCH v. 6.1, FastX v. , FastQC . Note that the following commands do not work with later versions of QIIME. Enable USEARCH to be executable with the following commands

```
chmod 755 usearch61
chmod 755 usearch
```

Make sure each of the programs can be accessed from any folder by typing the following commands (using the proper path for each program):

```
export PATH=${PATH}:/home/qiime/QWRAP
export PATH=${PATH}:/home/qiime/QWRAP/FastQC/
export PATH=${PATH}:/home/qiime/QWRAP/fastx/
export PATH=${PATH}:/home/qiime/usearch

source ~/.bashrc
```

Run a check to make sure they are all working correctly.

```
check_qwrap_plus.sh
```

They should all say Succeed. If not, try to chmod 755 them.

Sort your data: Find your data directory and put into a new folder where you will do all your analysis. It is easier to do this now. I named by new directory NEW_ANALYSIS but it shouldn't matter what you name yours. In the terminal, the rest of the scripts should be complete from inside this directory. Your sequencing data should be in an unzipped directory. However, the fastq files in each sample folder need to be zipped (fastq.gz).

To put the paired-end data into the correct format for QWRAP programs, you will need to put the first pass data (R1 files) into a directory called FORWARD and the reverse pass data (R2 files) into a directory called REVERSE. To do this directly, type the following commands:

```
mkdir FORWARD
mkdir REVERSE
shopt -s globstar
```

Then, to sort the files type this command, type

```
cp -r PH96Samples/**/*_R1* FORWARD/
cp -r PH96Samples/**/*_R2* REVERSE/
```

Go to the FORWARD and REVERSE folders outside the terminal and verify there are 96 files in each. Easiest to view if you click on View in the menu bar and click on List. If not, manually remove any files that are not supposed to be in there.

Merging the paired end reads into single end reads and trimming for quality

For Quality check before merging: run the program "quality_check_before.sh" with location of FWD reads and REV reads.

```
quality_check_before.sh FORWARD REVERSE
```

This script takes a while to complete. This creates two folders fastqc_beforef and fastqc_beforer with FASTQC results for the forward and reverse files and stats.

Merge the forward and reverse files: Run program: merge_reads_F_R.sh with parameters containing the location of forward and reverse folder. This creates a TEMP folder which has all the reads for the analysis.

```
merge_reads_F_R.sh FORWARD REVERSE
```

This script creates a TEMP folder with 192 items.

Prepare for merging reads: The program "prepare_merge_fastq.sh" requires the location of TEMP folder (containing both forward and reverse files) as a command line argument. Run the program as

```
prepare_merge_fastq.sh TEMP
```

You can edit the column 3 of the merged file in a text editor if required (especially if you want to rename the files. Make sure that the names in 3rd column do not include an underscore (_).

Merge the reads: This is done using program merge_fastq.sh. This script uses program USEARCH for merging reads. The program needs 5 parameters which includes the quality control parameters for merging.
    1)   Name of mapping file (Paired_Filelist.txt)
    2)   Length for trimming forward reads. Provide reads full length if no trimming is required.
    3)   Length for trimming reverse reads. Provide reads full length if no trimming is required.
    4)   Max mismatch allowed between forward and reverse reads when aligned.
    5)   Minimum overlap required between forward and reverse reads when aligned.

In these examples since we had reads of length 251 and decided not to trim them, we used the following parameters. This command makes ad Paired_Filelist.txt.

merge_fastq.sh Paired_Filelist FWD_TRIM REV_TRIM USEARCH_MAXDIFF USEARCH_MINOVERLAP

merge_fastq.sh Paired_Filelist.txt 250 250 15 50

This script takes a while to complete. This script creates all fasta files in the current directory after merging. It also stores the FASTQ files in a folder MERGED_FASTQ. Make sure you have this folder with 96 files in it before you go on to the next step.

Quality filtering after merging: run the program "quality_filter_single.sh" with location of merged FASTQ reads. The command line arguments are described above.

quality_filter_single.sh MERGED_FASTQ 250 80 20

This script takes a little while to complete. The program does the quality filtering and produces the filtered fastq files in a directory called "filtered_fastq". The program also creates the fasta file for all the samples in the current directory "ANALYSIS" which are used for subsequent analysis.

The program "quality_check_filterdata.sh" uses the directory filtered_fastq to generate the quality report for the filtered fastq files.

quality_check_filterdata.sh filtered_fastq

This script creates a folders fastqc_filterdata with FASTQC report for all files of the filtered dataset. Inside the folder, the HTML file "FASTQC_overview.html" is created which provides an combined overview of the quality statistics for all samples and also provide more detailed report for individual samples.

Processing the samples for 16S taxon identification and diversity

Customize the microbiome-workflow1.sh script according to whether you have 454 or Illumina reads. For this, display the script in gedit, find the RDP threshold and modify if needed. The threshold for 454 reads is the default and is 0.8. However, if you have Illumina reads, you will need to change this to 0.5 and save the script. Then run the following script in the terminal:

microbiome-workflow1.sh

This will generate five files in your NEW_ANALYSIS directory: seqs.fna, mapping.txt, sample_order.txt, config.log, and script.sh.

Define your samples: You will need to modify the mapping.txt to include some information about the samples. You will need to do this manually for all samples. Here is an example. You can see

I used Group 1 and Group 2 as the titles for these columns <u>which are separated with tabs (not spaces)</u>. In the first and third row below, I highlighted the stuff I manually entered for the visualization in this document (but not in the mapping.txt file). For the variables you can use categorical or quantitative data. In this example, samples were divided in Group 1 into antibiotic (Ab) or control (Ct). in Group 2, samples were divided by sampling site type (Ag – agricultural samples, Dr- drain samples, Nw- natural waters, Ww – wastewaters).

```
#SampleID       Group1          Group2
#Mapping file for the QIIME analysis
PH-01-Ag-TF5-AbS1L001R  Ab      Ag
PH-02-Ag-TF5-CS2L001R   Control         Ag
PH-03-Ag-RFF1-AbS3L001R         Ab      Ag
PH-04-Ag-RFF1-CS4L001R  Control         Ag
PH-05-Ag-MF2-AbS5L001R  Ab      Ag
PH-06-Ag-MF2-CS6L001R   Control         Ag
PH-07-Ag-WBF-AbS7L001R          Ab      Ag
PH-08-Ag-WBF-CS8L001R   Control         Ag
PH-09-Ag-WHF-AbS9L001R          Ab      Ag
PH-10-Ag-WHF-CS10L001R          Control         Ag
PH-11-Ag-SFR2-AbS11L001R        Ab      Ag
PH-12-Ag-SFR2-CS12L001R         Control         Ag
PH-13-Ag-SFR1-AbS13L001R        Ab      Ag
PH-14-Ag-SFR1-CS14L001R         Control         Ag
PH-15-Ag-SDDR2-AbS15L001R       Ab      Ag
PH-16-Ag-SDDR2-CS16L001R        Control         Ag
PH-17-Ag-SFR3-AbS17L001R        Ab      Ag
PH-18-Ag-SFR3-CS18L001R         Control         Ag
PH-19-Ag-DRAR-AbS19L001R        Ab      Ag
PH-20-Ag-DRAR-CS20L001R         Control         Ag
PH-21-Ag-GHF2-AbS21L001R        Ab      Ag
PH-22-Ag-GHF2-CS22L001R         Control         Ag
PH-23-Ag-RI1-AbS23L001R Ab      Ag
PH-24-Ag-RI1-CS24L001R  Control         Ag
PH-25-Ww-3OM-AbS25L001R         Ab      Ww
PH-26-Ww-3OM-CS26L001R          Control         Ww
PH-27-Ww-OXCM-AbS27L001R        Ab      Ww
```

Identify the taxa associated with each file: Execute the file "script.sh" as

`sh script.sh`

This script takes FOREVER (i.e. > 30 minutes; can take up to >60 minutes when using Illumina reads). This will execute all the commands present in the file and generate unfiltered/unrarified OTU table and the taxonomic charts. Please note here that the OTU table is not normalized for sample size differences and there is no filtering done at this step to remove any rare taxa. We call

this as original (all files include "_org" in their name). The taxa charts (taxa_summary_org), OUT table (otu_table_org.biom / otu_table_org.txt) are generated.

Steps displayed while running this command:
Chimera Filtering
OTU picking
Picking representative of OTUs
Assigning taxomony using RDP
Sorting OTU table
OTU table statistics
Summarizing taxa (Before filtering)
Converting BIOM file to TXT file
Creating Normalized OTU table

Folder/files created:
Folder: top_otu_taxa_org
Folder: taxa_summary_org
Folder: rdp_assigned_taxomony
Folder: uclust_picked_otus
Folder: usearch_checked_chimeras
File: normalized_otu_org.txt
File: otu_table_org.txt
File: otu_table_org.stats.txt
File: otu_table_org.biom
File: otu_table_unsorted_org.biom
File: seqs.fna_rep_set_org.fasta
File changed but not created?: seqs.fna

Remove chimera, filter and rarify your data, and identify the top taxa in each sample: The script requires one user defined parameter "sampling depth" (read depth). Since different samples may have different read depth, the read depth should be normalized across all samples. When a read depth is provided, a random sampling event is used to rarify the OTU table. If the sampling depth is 22986 (in our example dataset), you can run the script as

microbiome-workflow2.sh 22986

If no sampling depth is provided, the program will automatically calculate the minimum sampling depth from the file "otu_table_org.stats.txt". I don't think there is an output for this command and it should take ~1 second before the command prompt pulls up. File: script_adv.sh is created during this step.

Analyze the alpha diversity (within each sample) and beta diversity (between samples) and generate plots: The file "script_adv.sh" can be executed as

sh script_adv.sh

This script can also take a little bit of time (i.e. ~5-10 minutes). It is preparing several folders and graphs to compare alpha and beta diversity of the samples.

Steps involved in this script:
Rarefaction of OTU tables
Filtering OTUs at 0.0005% abundance
Summarizing taxa(filtered)
OTU table statistics
Converting BIOM file to TXT file
Creating Normalized OTU table

Folder/files created:
Folder: pynast_aligned
Folder: top_otu_taxa_fil
Folder: OTU_fil_Network
Folder: OTU_fil_Heatmap
Folder: taxa_summary_fil
Folder: alpha_rarefac
Folder: beta_div
Folder: beta_div_matrices
Folder: filtered_alignment
File: beta_params.txt
File: alpha_params.txt
File: phylogeny.tre
File: alpha_div.txt
File: normalized_otu_fil.txt
File: seqs.fna_rep_set_fil.txt
File: otu_table_fil.biom

Make your final report: Run the following program to generate the HTML report.

report_microbiome.sh

This creates an html file "microbiome_report.html" in the NEW_ANALYSIS directory and can be opened using any web browser. This file contains information about the original data (e.g. # of reads, quality of the reads), the data after QC filtering (# reads, quality of the reads), the OTU assignations for each sample, rarefaction tables, OTU charts, lists of the top 10, 25, and 100 taxa, and PcoA plots. It has also information regarding how to complete some statistical analysis though this is sorely lacking in my opinion. This document is the most important resource generated and it will be the basis of future analyses.

Folder/files created:
Folder: report_files
File: microbiome_report.html
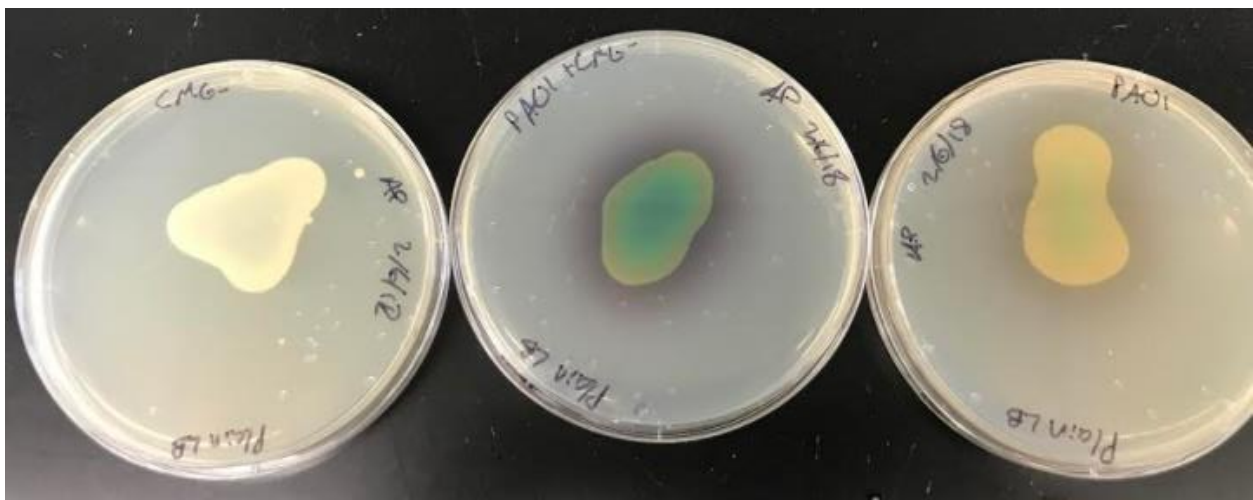
Supplementary Figure 1. From left, *Ochrobactrum* sp. CMG– only, mating mix of *Ochrobactrum* sp. CMG– and *P. aeruginosa* PAO1, and *P. aeruginosa* PAO1 only. All colonies were plated on plain LB.