

Copyright
by
Akshita Maheshkumar Mirani
2022

APPLICATION OF MALDI-TOF MS FOR MULTIPLE SOURCE TRACKING FROM
SEWAGE AND SEPTIC TANKS ALONG WITH IDENTIFICATION OF
ANTIBIOTIC RESISTANT PATHOGENIC ESCHERICHIA COLI

by

Akshita Maheshkumar Mirani, B.S.

THESIS

Presented to the Faculty of
The University of Houston-Clear Lake
In Partial Fulfillment
Of the Requirements
For the Degree

MASTER OF <SCIENCE/ARTS/BUSINESS ADMINISTRATION>
in Biotechnology

THE UNIVERSITY OF HOUSTON-CLEAR LAKE

MAY, 2022

APPLICATION OF MALDI-TOF MS FOR MULTIPLE SOURCE TRACKING FROM
SEWAGE AND SEPTIC TANKS ALONG WITH IDENTIFICATION OF
ANTIBIOTIC RESISTANT PATHOGENIC ESCHERICHIA COLI

by

Akshita Maheshkumar Mirani

APPROVED BY

Michael G. LaMontagne, Ph.D., Chair

Lory Z. Santiago-Vázquez, Ph.D., Committee Member

Jessica Labonté, Ph.D., Committee Member

Michael Allen, Ph.D., Committee Member

RECEIVED/APPROVED BY THE COLLEGE OF SCIENCE AND ENGINEERING:

David Garrison, PhD., Associate Dean

Miguel A. Gonzalez, PhD., Dean

Acknowledgements

First and foremost, I would like to express my thankfulness to University of Houston-Clear Lake (UHCL) for giving me an opportunity to perform research and for providing scholarships. Thank you to the UHCL library for their help in providing research papers as well as guidance for writing thesis.

I would like to give special thanks to my parents, brother, sister-in-law, and my entire family for all their support and understanding, without them it would not have indeed been possible.

Thank you Dr. LaMontagne, for believing in me and providing me a space to work on my own ideas. Working in your lab as a graduate student and a Teaching Assistant (TA) has made me more curious about microbiology. Your dedication towards research is, and will always be an inspiration for me. Thank you for being my thesis chair, it was my honor to be your student. I would like to thank my thesis committee member Dr. Santiago for your guidance and helping me with my thesis proposal. Thank you to my other committee members Dr. Allen and Dr. Labonté, your research has been a great guidance for me during my thesis.

I would like to acknowledge, Moody Gardens, Dallas Salmon WWTP, and Shamrock Septic from providing with seals, sewage, and septic tank samples respectively.

A special thanks to my lab partner Caroline Kmiecik, for helping me whenever I got stuck in my project, it was great teaming up with you. Thank you to all my other friends at UHCL.

ABSTRACT

APPLICATION OF MALDI-TOF MS FOR MULTIPLE SOURCE TRACKING FROM
SEWAGE AND SEPTIC TANKS ALONG WITH IDENTIFICATION OF
ANTIBIOTIC RESISTANT PATHOGENIC ESCHERICHIA COLI

Akshita Maheshkumar Mirani
University of Houston-Clear Lake, 2022

Thesis Chair: Michael LaMontagne, Ph.D.

The intensity of extreme flooding events, driven by tropical cyclones and sea-level rise, may increase dramatically this century. These extreme weather events can spread untreated sewage from wastewater treatment plants and onsite wastewater treatment systems (OWTS, septic tanks), which creates a possibility of outbreaks of water-borne diseases. Human waste represents a particular threat because it is laden with antibiotic-resistant bacteria. In particular, bacteria in waters that appear to be contaminated with human waste show a high level of resistance to the antibiotic polymyxin. To assess the risk of disease spread, managers need tools to track the source of this contamination. Escherichia coli (E. coli) is widely used as a fecal indicator bacteria (FIB). Elevated levels of this FIB suggest microbial contamination but not the source. Commonly used microbial source tracking tools, are time-consuming and expensive. Matrix-assisted laser

desorption/ionization-timeof-flight mass spectrometry (MALDI-TOF MS) is a time and cost-effective way to identify bacteria. This proteomics method can distinguish strains of bacterial species but has not yet been widely used for microbial source tracking. In this study, *E. coli* strains were isolated from wastewater treatment plants and a composite sample from OWTs. For comparison, a library of isolates was also generated from seal scat and dog feces. Isolates were then identified by MALDI-TOF MS, and cluster analysis was performed on mass spectra to determine if this technique can differentiate the sources of these FIB. To get knowledge on *E. coli* resistance towards antibiotics, a disc diffusion assay was implemented to screen representative isolates for sensitivity to broad spectrum antibiotics and polymyxin B. Colistin resistance of representative isolates and likely FIB sources (sewage, OWTs, seal scat, and dog feces) was checked by PCR using primers pairs specific for the genes *mcr-1*, *mcr-2*, and *mcr-3*. MALDI-TOF MS distinguished *E. coli* strains isolated from sewage and OWTs from *E. coli* isolated from animal sources. Antibiotic resistance assays indicated that *E. coli* strains isolated from all sources were resistant to ampicillin, streptomycin, and gentamicin. Bacteria isolated from sewage and OWTs showed resistance to colistin and polymyxin B and DNA extracted from a sewage sample tested positive for colistin resistance with primers specific for *mcr-3*. The results suggests that MALDI-TOF MS could be applied to track the source of fecal contamination of waterways. This could improve risk assessment and point to mitigation strategies. Preliminary results also suggest that polymyxin resistant *E. coli* strains are common in sewage. Screening for these antibiotic resistant genes could indicate the presence of human waste.

TABLE OF CONTENTS

List of Tables	ix
List of Figures	x
CHAPTER I: INTRODUCTION.....	1
Concern for Increasing Flooding events	1
Presence of FIB.....	3
MALDI-TOF MS.....	4
Antibiotic Resistance	5
Colistin And Polymyxin B Resistance.....	5
CHAPTER II: METHODOLOGY	7
Expected Taxa from Meta microbiome analysis	7
Sewage, Septic Tank, and Animal Scat Sampling.....	9
Culturing	11
MALDI- TOF MS.....	11
DNA Extraction, Concentration, and Quantification.....	12
Environmental sample DNA Extraction	12
E. coli isolate DNA Extraction	13
DNA Concentration and Quantification.	13
Antibiotic Resistance	14
Kirby-Bauer	14
Colistin Resistance- PCR.....	15
CHAPTER III: RESULTS.....	17
Sampling and Culturing	17
Composition of Sewage by MADLI-TOF MS	17
E. coli screening and isolation	19
E. coli Identification by MALDI-TOF MS.....	21
DNA Extraction	22
Isolate DNA extraction	22
DNA Concentration	26
Total DNA Extraction.....	29
DNA Quantification.....	31
Antibiotic Resistance	33
Kirby Bauer for Representative Microorganism.....	33
E. coli Antibiotic Resistance Broad Spectrum.....	33
E. coli Antibiotic Resistance Narrow Spectrum	34
Colistin Resistance Evaluation by PCR.....	39

CHAPTER IV: DISCUSSION	44
E. coli Isolation	44
MALDI-TOF MS.....	44
Antibiotic Resistance	47
Limitations	47
CHAPTER V: CONCLUSION AND FUTURE DIRECTION.....	49
REFERENCES	50
APPENDIX.....	59
Bioinformatics Script:	59
G-Block sequence:	60
iTecan plate reader parameters:	61
MTU table from R-script combined with Bruker ID and score:.....	62
Kirby-Bauer results for sewage isolates:	63

LIST OF TABLES

Table 1. Meta-microbiome analysis of sewage and OWTS.	7
Table 2. DNA standards Fluorescence after running through AccuBlue® quantification	14
Table 3. Antibiotics used for Kirby- Bauer with their concentration	15
Table 4. PCR primers and primer sequences	16
Table 5. PCR master mix components.....	16
Table 6. Cycling conditions for PCR.....	16
Table 7. Nanodrop readings for DNA extraction by for each isolate	23
Table 8. Samples after DNA extraction added in each lane of agarose gel with 5 µl volume in each well.....	24
Table 9. Nanodrop readings after DNA purification with magnetic beads	27
Table 10. Samples after DNA concentration added in each lane of agarose gel with 5 µl added to each well.	28
Table 11. Meta samples extraction added in each lane of agarose gel with 5 µl of volume in each well.	30
Table 12. Nanodrop readings for total DNA extracts from meta-samples	31
Table 13. DNA quantification by AccuBlue®, shows fluorescence value by 96- well plate reader along with the samples nanodrop readings. Yellow highlights were the isolates used for antibiotic resistance assays.....	32
Table 14. Antibiotic resistance data for E. coli from multiple sources with MTU and the zone of inhibition	34
Table 14 cont. Kirby Bauer protocol for Polymyxin B resistance on E. coli isolates and for direct samples from seal and sewage.....	36
Table 15. Samples after DNA amplification with 16S primer added in each lane of agarose gel with 5 µl added to each well.	40
Table 16. Samples after DNA amplification with mcr-1 primer added in each lane of agarose gel with 5 µl added to each well.....	41
Table 17. Samples after DNA amplification with mcr-2 primer added in each lane of agarose gel with 5 µl added to each well.....	42
Table 18. Samples after DNA amplification with mcr-3 primer added in each lane of agarose gel with 5 µl added to each well.....	43

LIST OF FIGURES

Figure 1. Nonmetric multidimensional scaling analysis (NMDS) of microbial community structure for samples collected from Clear Lake (TX) before and after Hurricane Harvey (HH).	2
Figure 2. Top 20 bacterial taxa data for 16S rRNA gene sequences for sewage (out719) and OWTS (out717) fastq files were processed with the DADA2 pipeline, showing majority in both samples to be Gammaproteobacteria class.	8
Figure 3. Bacterial taxonomic composition obtained for sewage samples shotgun data processed with ccMetagen.	9
Figure 4. WWTP sample collection facility; Dallas Salmon WWTP.....	10
Figure 5. WWTP sample collection facility anonymous sewage treatment plant	10
Figure 6. DNA standards linear graph made in SigmaPlot 14.5; the graph shows R ² value along with linear graph equation.	14
Figure 7. Sewage samples directly spread on A. MacConkey Agar and B. EMB Agar.....	17
Figure 8. Score values obtained by comparing sample spectra with reference known microbes	18
Figure 9. Dendrogram from sewage samples showing the sewage composition.	18
Figure 10. Taxonomic composition of the sewage isolated identified by MALDI-TOF.	19
Figure 11. 96 well plate wells glowing under UV light indicate positive for presence of E. coli.....	20
Figure 12. Presence of E. coli on A. EMB plate, B. MacConkey plate.....	21
Figure 13. Restreaked isolated colonies on EMB.....	21
Figure 14. Dendrogram for E. coli from multiple sources; Sewage, Septic tanks, Dog feces and Seal scat, with bootstrap values in red numbers.	22
Figure 15. Gel 1, DNA from isolates (1% Agarose gel, in 1X TBE, Gel Red, 150V, 45 mins, 1kb plus DNA ladder).	25
Figure 16. Gel 2, DNA from isolates (1% Agarose gel, in 1X TBE, Gel Red, 150V, 45 mins, 1kb plus DNA ladder).	26
Figure 17. DNA from isolates after concentration by magnetic beads (1% Agarose gel, in 1X TBE, Gel Red, 150V, 45 mins with 1kb plus DNA ladder).	29
Figure 18. Total DNA extraction from meta-samples (1% agarose gel made in 1X TBE, Gel Red, 150V, 45 mins, 1Kb DNA Ladder).....	30

Figure 19. Comparison between concentration of DNA from Nanodrop (x-axis) and concentration of DNA from AccuBlue® (y-axis) made in SigmaPlot 14.5	33
Figure 20. Resistance for E. coli in conjunction with different antibiotics.	36
Figure 21. Antibiotic resistance plates with 6 different antibiotic discs in each plate showing zone of clearance for susceptibility.	37
Figure 22. Polymyxin B Antibiotic Resistance by Kirby-Bauer on E. coli and for direct samples from seal and sewage.	38
Figure 23. PCR Amplification with 16S rRNA gene primers (1% Agarose gel with 1X TBE, Gel Red, 150V, 45 mins with 1 kb DNA plus Ladder), shows amplification for all isolates.....	40
Figure 24. PCR Amplification with mcr-1 primers (1% Agarose gel with 1X TBE, Gel Red, 150V, 45 mins with 1 kb DNA plus Ladder), shows unexpected bands for sewage (lane 6) and SR27 (lane 4) samples, when compared with positive control (lane 1).	41
Figure 25. PCR Amplification with mcr-2 primers (1% Agarose gel with 1X TBE, Gel Red, 150V, 45 mins with 1 kb DNA plus Ladder), gel shows expected results with low molecular bands for all samples.....	42
Figure 26. PCR Amplification with mcr-3 primers (1% Agarose gel with 1X TBE, Gel Red, 150V, 45 mins with 1 kb DNA plus Ladder), gel shows expected results for sewage sample (lane 6) as compared to G-block (lane 1).	43
Figure 27: Hierarchical clustering based on core genome multilocus sequence typing.	46

CHAPTER I: INTRODUCTION

Concern for Increasing Flooding events

Climate change will increase the intensity of tropical storms (Knuston et al. 2020) and has increased the risk of flooding in coastal regions by approximately 35% (Hettiarachchi et al. 2018). Flooding increases the risk of pathogen transmission and is a major issue for cities along the coast of the Gulf of Mexico (Sericano et al. 1994), especially Houston (Pingfeng et al. 2018). For example, in 2017 Hurricane Harvey flooded 25-30% of the city of Houston, resulting in the spread of untreated sewage from wastewater treatment plants (Stuckey, 2017). Targetted metagenomic analysis suggests this event contaminated tributaries to Galveston Bay with human waste (Fig. 1). Flooding from Hurricane Laura in 2020 spread untreated sewage in the Houston-Galveston area (Rasha et al. 2020). These floods carry pathogens that may drive outbreaks of water-borne infectious diseases and create new threats, including antibiotic resistant bacteria (Garner et al. 2017).

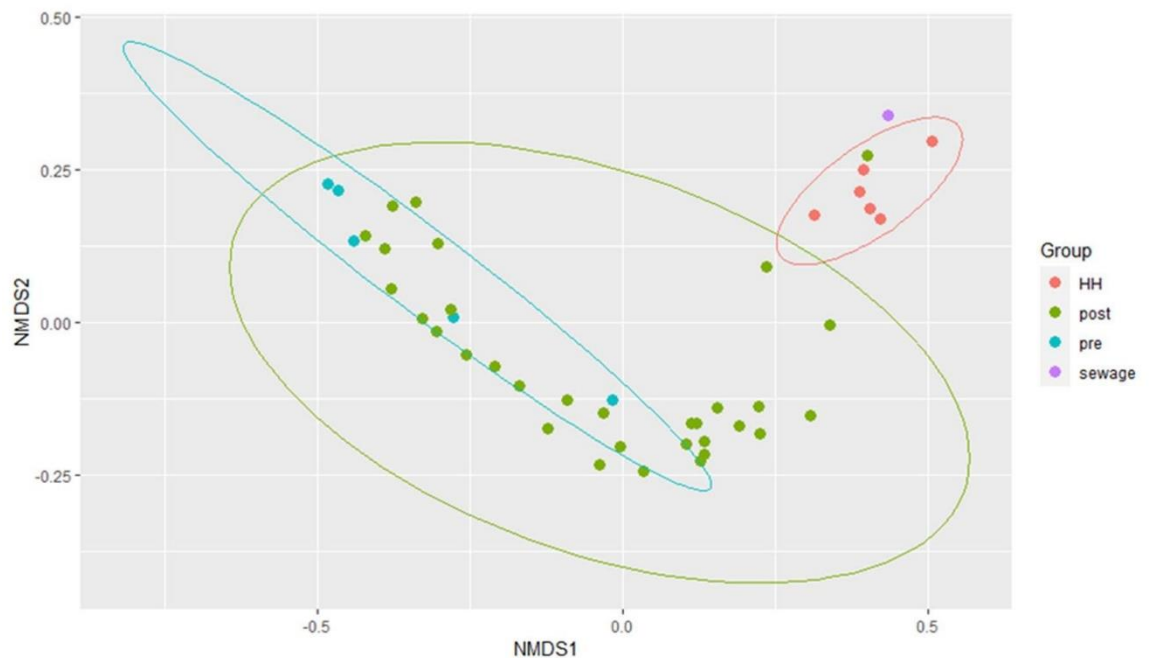


Figure 1. Nonmetric multidimensional scaling analysis (NMDS) of microbial community structure for samples collected from Clear Lake (TX) before and after Hurricane Harvey (HH). Bacterial communities in floodwater (HH (red)) showed similarity to communities in a water sample spiked with raw sewage (purple). Figure reproduced with permission from LaMontagne et al. 2022.

Watershed coupled to Western Galveston Bay contain thousands of onsite wastewater treatments systems (OWTS), which can be an important nonpoint source of fecal contamination (Withers. 2014). This contamination threatens the economic viability of communities along the bay and, during extreme weather events, the health of residents exposed to contaminated floodwaters (Yang. 2021). However, this public health threat is not well studied. In particular, relatively few studies have described bacterial communities in OWTSs.

Penicillin was discovered by Sir Alexander Fleming in 1928 and first became commercially available to the general public in 1945. The first case of penicillin-resistance was observed shortly after in 1947 (Lobanovska et al. 2017). Many of the antibiotics currently used were produced during the “golden age” of antibiotic production

in the 1950s and 1960s (Davies. 2006). Since then, the use of antibiotics has increased dramatically and The Center of Disease Control (CDC) has classified antibiotic resistant bacteria (ARB) as a severe threat (CDC. 2019). Additionally, the CDC has classified a few strains of bacteria as a severe concern, including colistin-resistant *E. coli* and vancomycin-resistant Enterococci; both of which have been found to be present in sewage (Eramo et al. 2017).

Fecal indicative bacteria (FIB) are found in the gut of mammals and are almost universally used to assess water quality. FIB counts provide useful information to stakeholders; however, deeper studies of the prevalence and persistence of the ARB and antibiotic resistant genes (ARG) in FIB collected during water quality monitoring are needed to further assess public health risks. FIB frequently exchange ARG through horizontal gene transfer (Granado et al. 2019, Pandey et al. 2014, Ventola 2015). Floodwater generated by extreme weather events could spread these ARB and ARG through aquatic systems (Kaufmann et al. 2016). However, the connection between the increase in floods and spreading of ARB and ARG has yet to be well characterized. To track the spread of these emerging threats, routine and consistently monitored stations need to be implemented as well as genetic analysis of ARB (Gong et al. 2019, Kawecki et al. 2017).

Presence of FIB

Fecal contamination in freshwater can be determined by elevated counts of *E. coli* as recommended by the U.S Environmental Protection Agency (EPA). Elevated counts of this FIB suggest an increased risk of water-borne diseases. However, the increase in counts do not indicate the source. Therefore, source tracking protocols are implemented.

Bacterial source tracking methods can be divided into culture-independent and culture-dependent approaches (Field. 2002). Culture-independent methods include targeted and shotgun metagenomics (Raza 2021, Gomi 2014), and most commonly,

qPCR with source specific primers (Aridi et al. 2020). Culture-dependent methods generally involve creating libraries of FIB from sources and samples and then determining the frequency of FIB associated with different sources. This laborious bacterial source tracking approach has particular value for managers because they can identify the source or sources directly responsible for elevated FIB counts. This informs mitigation strategies and provides legal evidence for enforcement. Culture-independent methods include targeted and shotgun metagenomics

Culture-dependent protocols of bacterial source tracking, including antibiotic resistant patterns and ribotyping, can give ambiguous results. (Stoppe et al. 2014). Whole genome sequencing (WGS) has emerged as a powerful tool for tracking infectious disease outbreaks (Brown et al. 2021, Nouws et al. 2021) but WGS is expensive and time consuming. Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) on the other hand is a fast and inexpensive method of bacterial identification, and has resolution comparable to WGS (Marzhari 2021).

MALDI-TOF MS

Application of MALDI-TOF MS for microbial identification involves ionization of protein extracts with a matrix and then blasting the sample with short laser pulses. The resulting charged peptide fragments are accelerated by electromagnetic fields in a vacuum tube (Wieser et al. 2012). The time fragments take to reach the detector is proportional to their mass and the degree of ionization. This creates a mass spectrum fingerprint, which can be used to rapidly and reliably identify microbes (Maier et al. 2006). This high throughput method can provide results for hundreds of isolates in an hour. Commercially available MALDI-TOF systems, use proprietary software to compare spectra from isolates against a reference database for microbial identification. The popular Bruker MALDI Biotyping system (Bruker Diagnostics) gives each fingerprint a reliability score (1.5-3.0) based on matches to reference spectra. A score

above 2.3 is considered species-level identification. MALDI-TOF MS can also discriminate strains of bacterial species including antibiotic resistance strains of pathogens (Singhal et al. 2015, Motlagh and Yang 2019).

Antibiotic Resistance

More than 2.8 million antibiotic-resistant infections occur in the U.S. each year (CDC, 2021) and ARB are common in human waste. For example, 92 *E. coli* strains isolated from children, 40% of the isolates were ampicillin resistant, 25% were tetracycline resistant, and 26% were streptomycin resistant (Barreto et al. 2009).

To determine resistance of bacterial strains, antibiotic representatives from each class should be considered. These include, chloramphenicol, ampicillin, tetracycline, gentamicin, erythromycin, and streptomycin.

Colistin And Polymyxin B Resistance

Resistance to polymyxin is an emerging threat (Li et al. 2019) and floodwaters generated by Hurricane Harvey appear enriched in genes associated with resistance to polymyxin (LaMontagne et al. 2022). Polymyxin class antibiotics include polymyxin E (colistin) and polymyxin B. These antimicrobials are one of the last remaining treatments for multidrug-resistant Gram-negative bacteria, especially Enterobacteriaceae (Li et al. 2019). Colistin antibiotic is a circular peptide which disrupts the outer membrane of Gram-negative bacteria (Gogry et al. 2021). Bacteria acquire resistance against this class of antibiotics by mutation in genes that code for lipopolysaccharides (Li et al. 2020). Resistance to colistin is common in bacteria isolated from livestock and humans (Olaitan et al. 2014) and is associated with virulent strains. For example, colistin-resistant *E. coli* strain NC101 releases colibactin, a genotoxic metabolite that can break DNA and promote colorectal cancer (Silpe et al. 2022). Resistance to colistin is encoded in the genes *mcr*-1, 2 and 3 gene (Chalmers et al. 2018). These genes can be rapidly and sensitively detected by PCR amplification (Li et al. 2017, Reinthaler et al. 2003).

In this thesis project, I assessed diversity of *E. coli* strains isolated from sewage and OWTS samples. To provide preliminary data on the microbial communities in these water treatment systems, I did a meta-microbiome analysis of publicly available data. I then evaluated MALDI-TOF MS as a method of tracking these wastewater sources of *E. coli*. Seal scat and dog feces were used for comparison. To track the spread of ARB from these sources, antibiotic resistance for *E. coli* isolates was also evaluated and the prevalence of ARG in waste and fecal sources was also evaluated.

CHAPTER II: METHODOLOGY

Expected Taxa from Meta microbiome analysis

To compare bacterial communities in between sewage and OWTS, I analyzed sequences obtained from the European Nucleotide Archive (ENA) and the NCBI Short Read Archive (Table 1). The 16S rRNA gene data, the reads were subset by seqtk and processed by DADA2 pipeline (Appendix), which gave the top 20 bacterial taxa (Fig. 2).

Shotgun data were also subset by seqtk but analyzed with Aldex2 pipeline (Appendix). Shotgun data were viewed with the online web-browser ccMetagen to obtain a composition figure (Fig. 3). The majority of the bacteria found were member of the

Gammaproteobacteria class, which includes the Enterobacteriaceae family, and the Bacteroidia class.

Table 1. Meta-microbiome analysis of sewage and OWTS.

Metadata	SRA	Platform	Strategy	Location
OWTS	SRR11314289	Illumina Truseq	16S V4	United states
Sewage Water	SRR8573788	Illumina HiSeq	Shotgun metagenomics	Montevideo (capital of Uruguay)
Wastewater	PRJNA264400	Illumina MiSeq	16S V4-V5	71U.S. cities
Wastewater	PRJNA505617	Illumina	16S V4-V6	Hong Kong

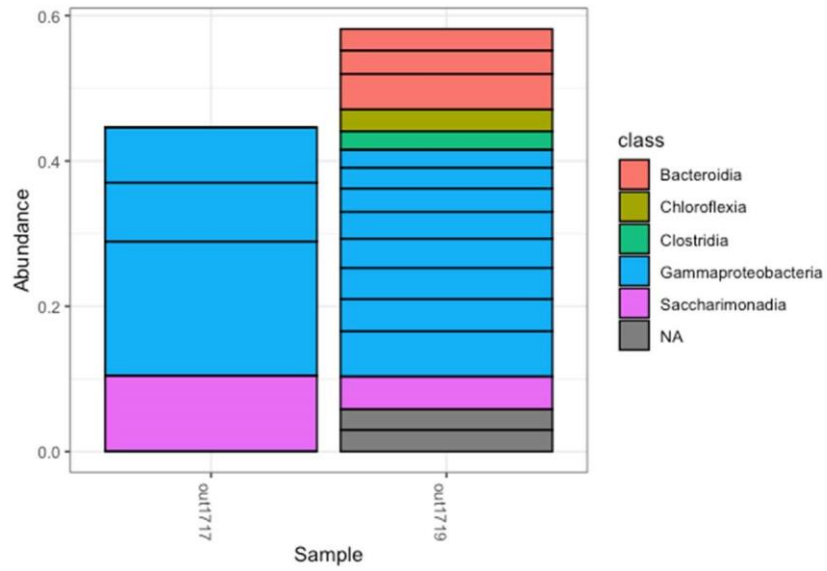


Figure 2. Top 20 bacterial taxa data for 16S rRNA gene sequences for sewage (out719) and OWTS (out717) fastq files were processed with the DADA2 pipeline, showing majority in both samples to be Gammaproteobacteria class.

2021 and dog feces were collected from two mixed-breed one-year old companion animals on November 23rd 2021.



Figure 4. WWTP sample collection facility; Dallas Salmon WWTP

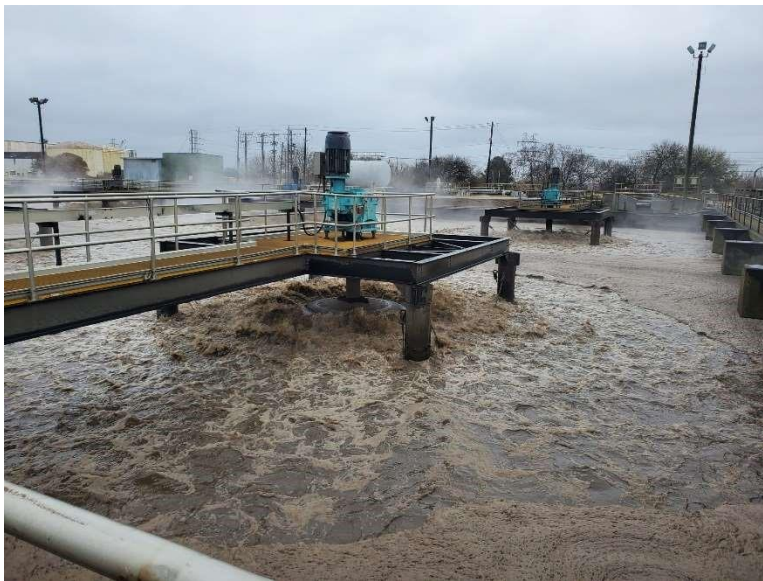


Figure 5. WWTP sample collection facility anonymous sewage treatment plant

All fecal and waste samples were collected in sterile 50 mL falcon tubes (Stellar Scientific). The tubes were transported back to the laboratory on wet ice. Sewage and OWTS samples (25 mL) were directly used without any processing. Seal scat and

dog feces were suspended in 25 mL 1X PBS. The biomass was concentrated from samples by centrifugation at 8,000 xg for 10 mins at 4°C. The supernatant was discarded in 10% bleach solution. Pellet obtained was resuspended in 2mL DESS solution (DESS was prepared by 20% DMSO, 500 mM EDTA, and 5M NaCl). DESS was used because DMSO helps in cell penetration of EDTA and NaCl, EDTA ceases nuclease activity, and NaCl ceases enzymatic activity. The DESS suspension was transferred to 4 x 1.5 mL sterile microcentrifuge tubes. The tubes were centrifuged at 16,000 xg for 2 mins, the supernatant was discarded, and the pellets were stored at -80°C.

Culturing

Differential and selective media were used to isolate the Gram-negative bacteria. One library was generated by diluting samples with in 1X PBS to 10^{-4} to 10^{-6} dilutions. 100 µl of each diluted sewage sample was directly spread on differential media (BB1 Levin Eosin Methylene Blue (EMB, BD-211221) and Difco MacConkey agar (BD-212123)). As the above approach yielded few E. coli isolates, samples dilutions were also mixed with equal volumes of 2X concentrations of primary media and were then dispersed into 96 well plates. Primary media for E. coli was A1 broth (Sigma 17112) supplemented with 20 mg/mL MUG working solution, which was prepared by diluting 100 mg MUG (Cayman 17205) with DMF (Sigma D-4551). Positive wells after checking under UV light were further plated on secondary media EMB and MacConkey. Isolates which showed positive wells along with growth on secondary media were re-streaked on to Tryptic Soy Agar (TSA, TEKNOVA- T0401) plates.

MALDI- TOF MS

Isolates from TSA were identified using MALDI-TOF MS with a 3-step process. Individual colonies from TSA plates were selected, processed by ethanol treatment, extracted with formic acid, and spotted on steel targets following protocols recommended by Bruker Scientific (Billerica MA, USA). Overnight bacterial cultures were suspended

in 300 µl of HPLC-grade water and 900 µl of HPLC-grade ethanol. The suspensions were stored at 4°C for up to 7 days. Suspensions were centrifuged at 14,500 xg for 2 mins to recover bacterial cells. 50 µl of each 70% formic acid and 100% acetonitrile were added, followed by centrifugation. The supernatant resulted in protein extracts. 384 polished/ 96 polished MALDI target steel plates were cleaned with freshly prepared 70% ethanol and 80% Trifluoroacetic Acid (TFA) as recommended by Bruker. Targets were spotted with 1 µl of extracts. Controls spotted received either 1 µl Bacterial Test Standard (BTS) or 1 µl of matrix solution. Targets were allowed to air dry and then shipped overnight with an ice pack to the Proteomics and Mass Spectrometry Core Facility at the Huck Institute (The Pennsylvania State University, University Park, PA, USA).

Mass spectra obtained after mass calibration with BTS were analyzed using an R script (R markdown file, see Appendix) following LaMontagne et al (2021). This R script uses the packages MALDIQuant, MALDIquant foreign, Rweka, pvclust, ggplot2, iNEXT, and philentropy. Multiple parameters were considered in the R script, including smoothing, baseline removal, alignments, signal to noise ratio (SNR) and peak detection. The script used two optimization loops. The first loop, optimized peaks and reports Jaccard coefficients of technical replicates, and the second loops was used for quality control and to optimize cosine similarity values calculated from peak heights for technical replicates. These values provide thresholds to define MALDI-TOF MS taxonomic units (MTUs), which are analogous to OTUs in metagenomic analysis.

DNA Extraction, Concentration, and Quantification

Environmental sample DNA Extraction

Total DNA was extracted from the samples stored in DESS at -80°C with the Macherey-Nagel NucleoSpin DNA stool kit (740472), with the following modifications, the addition of 850 µl of ST1 (from kit) was directly added to the pellets and this suspension was transferred into bead tubes. Also, DNA was eluted twice.

E. coli isolate DNA Extraction

For DNA extraction from isolates, two representative from each MTUs were selected from extraction. MTUs were defined from MALDI-TOF data following LaMontagne et al, (2021). For pure colonies, bacteria were restreaked on to fresh TSA plates and cultured in overnight shaker at 35°C in TSB medium. Overnight cultures were used for DNA extraction by Lucigen (Masterpure, MC85200) as per the manufacturer's instructions (Lucigen, MasterPure™ Complete DNA and RNA Purification Kit. Biosearch Technologies), with the following modifications. First, the initial heating was done at 65°C for 15 mins with shaking at 300 rpms. Second, the volume of RNase was increased to 2 µl of 100 ng/µl of RNase, which resulted in less RNA contamination. Finally, two elutions were performed to maximize recovery.

DNA Concentration and Quantification.

To remove small fragments and impurities, magnetic beads were used (Omega Mag-bind M1378-00). The ratio of beads to sample was 0.5X, with 10 µl of beads and 5 µl of sample. The rest of the protocol was followed using the short fragment removal (SFR) method (Maggie Weitzman, GC3F). After each step, DNA concentration was checked using two methods: Nanodrop (Thermo Scientific 2000) and agarose gel electrophoresis. After final DNA concentration, each sample was quantified using AccuBlue® (Biotium, AccuBlue® Broad Range dsDNA Quantitation Kit 31007-T). Protocol was followed as per the manual of instructions. Readings were taken with 96 well plate reader Tecan-i-control software, and parameters for the plate reader are in Appendix.

The concentrated DNA was quantified with the AccuBlue® high molecular DNA quantification kit. The standard obtained (Table 2) was plotted in SigmaPlot 14.5 application (Fig. 6). These standards were used to quantify the unknown sample concentration. A reliable R^2 value of 0.98 was achieved.

Table 2. DNA standards Fluorescence after running through AccuBlue® quantification

DNA	Fluorescence	DNA conc
2	15235	6.585559
6.25	15655	9.362602
12.5	15713	9.746099
25	16491	14.89024
50	21553	48.36022
100	31424	113.6273
150	35895	143.1896

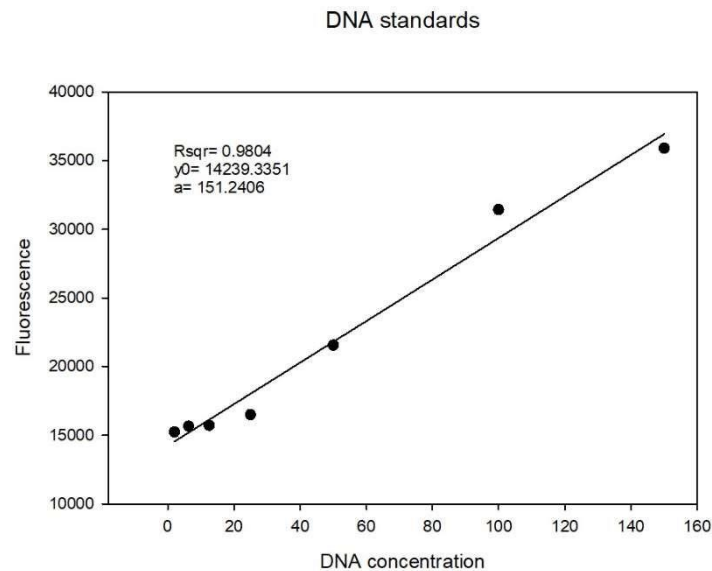


Figure 6. DNA standards linear graph made in SigmaPlot 14.5; the graph shows R^2 value along with linear graph equation.

Antibiotic Resistance

Kirby-Bauer

To screen for antibiotic resistance, representative MTUs were selected from the library generated from sewage samples and *E. coli* strains from different sources.

Bacterial suspensions from fresh overnight cultures in tryptic soy broth were diluted in 1X PBS and measured with a spectrophotometer to obtain an $A_{600\text{ nm}}$ of 0.07-0.08. From this

a dilution with approximately 10^5 CFU ml⁻¹ solution, 100 µl was spread on Difco-Mueller Hinton (BD- 225250) agar plates. Antibiotic discs were placed using a multidisc dispenser. Plates were incubated at 37°C for 24 hours. Results were recorded by measuring the zone of inhibition for each disc.

Table 3. Antibiotics used for Kirby- Bauer with their concentration

Antibiotic Disc	Concentration	Reference
Chloramphenicol	30 µg	BD 230733
Streptomycin	10 µg	BD 230942
Gentamycin	10 µg	BD 231227
Tetracycline	30 µg	BD 230998
Ampicillin	10 µg	Gibco™ 11593027
Erythromycin	15 µg	BD 230793
Polymyxin B	300 µg	OXOID CT0044B

Colistin Resistance- PCR

Primers mcr-1, mcr-2, and mcr-3 were used to check colistin resistance in *E. coli* (Table 4). For a positive amplification control, a G-Block that contained sequences for mcr1, mcr-2 and mcr-3 concatenated together was synthesized (Appendix). The compatibility for loop formation was checked on IDT website, using gBlocks® gene fragments entry tool. 2X Platinum II PCR master mix was used for the reaction. The cycling conditions were determined based on the melting temperature for primers and extension temperature for the Taq DNA polymerase (Table 6).

Table 4. PCR primers and primer sequences

Primer name	Sequence
mcr-1 forward mcr-1 reverse	5'AAAGACGCGGTACAAGCAAC 3'GCTGAACATACACGGCACAG
mcr-2 forward mcr-2 reverse	5'CGACCAAGCCGAGTCTAAGG 3' CAACTGCGACCAACACACTT
mcr-3 forward mcr-3 reverse	5' ACCTCCAGCGTGAGATTGTTCCA 3' GCGGTTTCACCAACGACCAGAA

Table 5. PCR master mix components

Component reaction	Final concentration	Volume
Platinum II PCR	1X	63 µl
Forward Primer (10uM)	500 nM	0.5 µl
Reverse Primer (10uM)	500 nM	0.5 µl
Nuclease free water		53 µl

Table 6. Cycling conditions for PCR

Step	Temperature	Time
Initial Denaturation	94°C	2 min
32 cycles		
Denaturation	94°C	15 secs
Annealing	60°C	15 secs
Extension	68°C	24 secs
Final Elongation	68°C	5 mins
Hold	4°C	∞

CHAPTER III:

RESULTS

Sampling and Culturing

Direct culturing of sewage from samples on differential and selective media yielded multiple colony morphologies (Fig. 7). MacConkey showed pink, white, orange-colored colonies, and EMB showed metallic sheen, orange, brown, transparent colored colonies.

Apparently, fast growing species took over the plate (Fig. 7), Dilutions were tried until 1:100 (Fig. 7 A) and still were obtaining same results as lower dilutions. And therefore, we also used A1 as a primary media for culturing *E. coli*. A1 media is a selective and differential media for specifically *E. coli*.

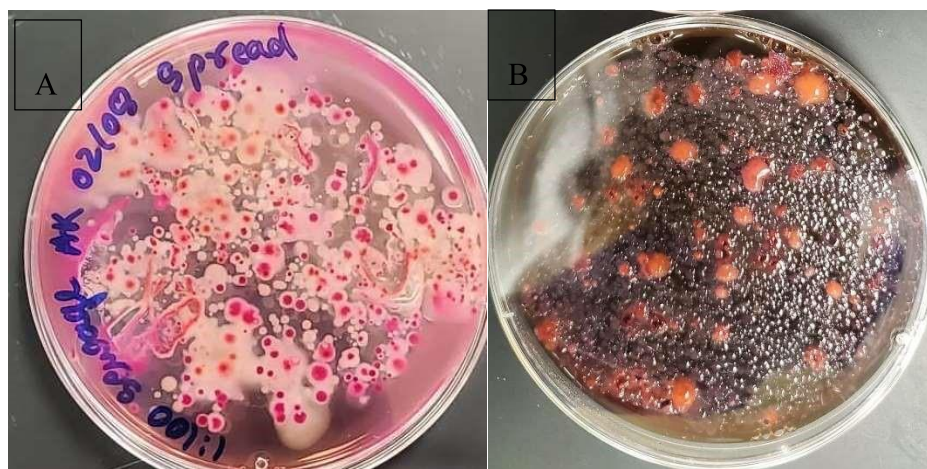


Figure 7. Sewage samples directly spread on A. MacConkey Agar and B. EMB Agar.

Composition of Sewage by MADLI-TOF MS

A total of 103 bacteria were isolated from sewage and OWTS samples directly plated on primary media. Most of these were identified at the species level by MALDITOF MS analysis, with a score of 2.0 and above (Fig. 8). These isolates were clustered into coherent MTUs. A majority of the isolates were identified as *Pseudomonas* (Fig. 10). There were few isolates which were in just 1% (others) *Shewanella* sp,

Kluyvera sp, Wautersiella sp, Tsukamurella sp, and Acinetobacter sp. The other (NA), could not be identified by reference genomes available. This culture-dependent library showed overlap with the structure of bacterial communities in wastewater, as assessed by meta-microbiome analysis. Analysis of publicly available 16S rRNA gene sequences (Fig. 2) and shotgun metagenomic data (Fig. 3), suggested that the majority of the bacteria wastewater classify as Gamma-proteobacteria, which included Enterobacteriaceae, and Bacteroidia.

Meaning of Score Values

Range	Description	Symbols	Color
2.300 ... 3.000	highly probable species identification	(+++)	green
2.000 ... 2.299	secure genus identification, probable species identification	(++)	green
1.700 ... 1.999	probable genus identification	(+)	yellow
0.000 ... 1.699	not reliable identification	(-)	red

Figure 8. Score values obtained by comparing sample spectra with reference known microbes

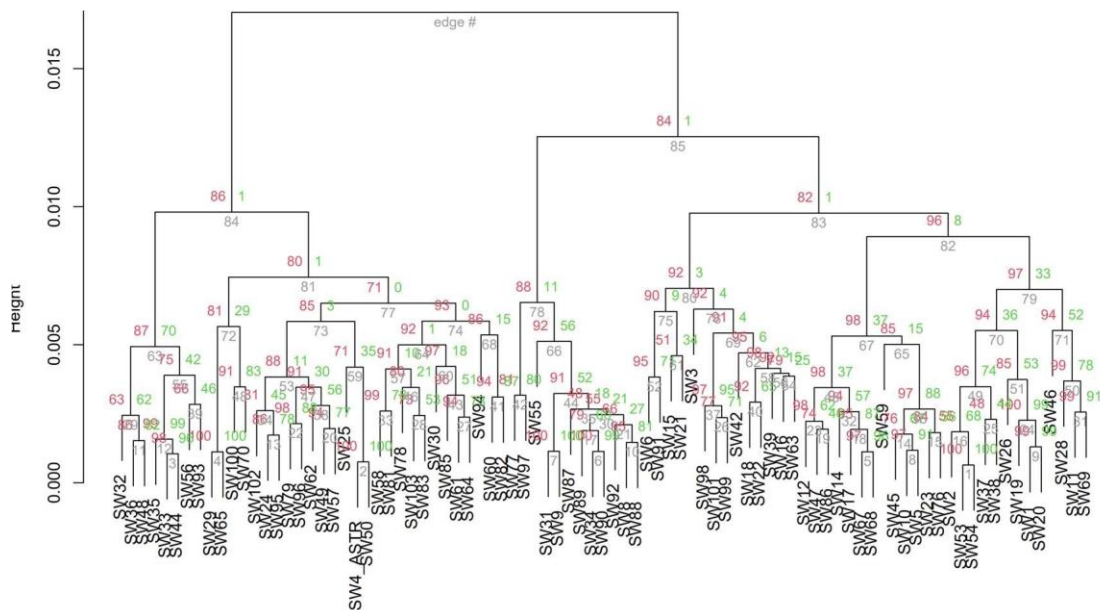


Figure 9. Dendrogram from sewage samples showing the sewage composition.

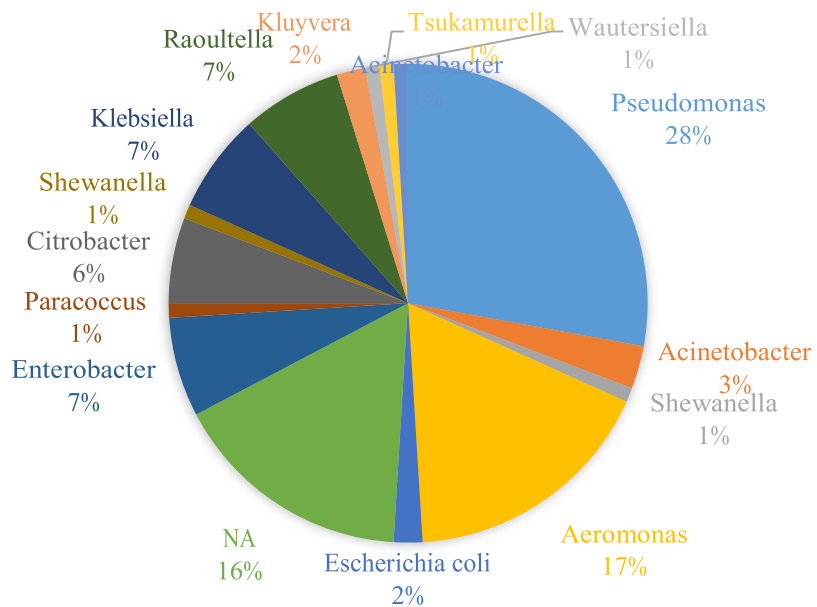


Figure 10. Taxonomic composition of the sewage isolated identified by MALDI- TOF.

E. coli screening and isolation

Sample dilution from 10^{-4} to 10^{-6} from sewage, septic tank along with seal scat and dog feces were inoculated into 96-well plates. After 24 hours, the plates were observed under UV light. Positive wells glowed under UV, due to interaction with MUG compound in media (Fig. 11).

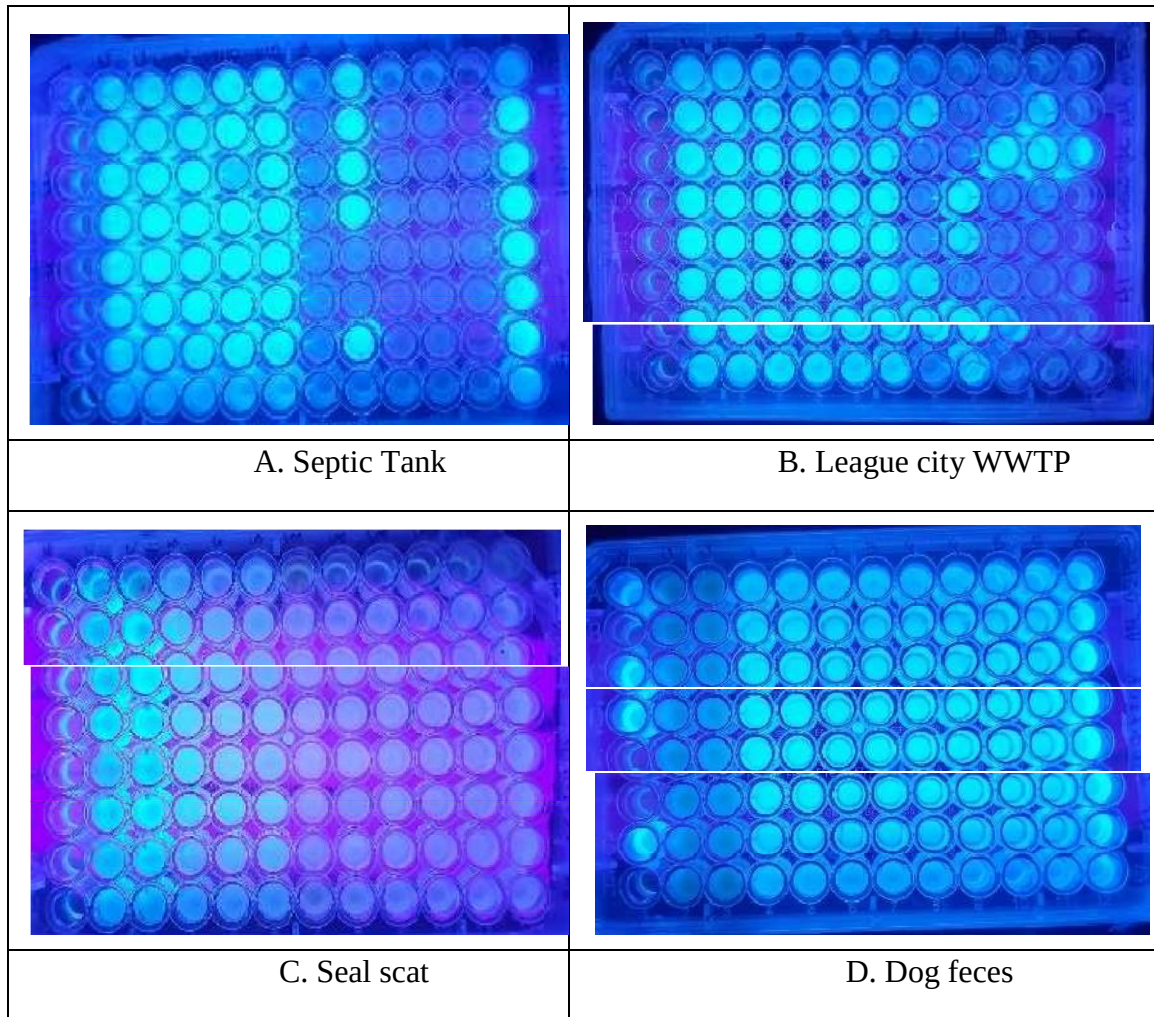


Figure 11. 96 well plate wells glowing under UV light indicate positive for presence of E. coli

Positive wells from A1 media were streaked onto secondary media (EMB and MacConkey, Fig. 12). Rapid lactose producers on EMB plate showed a metallic sheen which is consistent with E. coli and MacConkey plate showed dark pink color colonies, which is also consistent with E. coli. Isolated colonies from these plates were restreaked again onto the EMB for confirmation (Fig. 13).

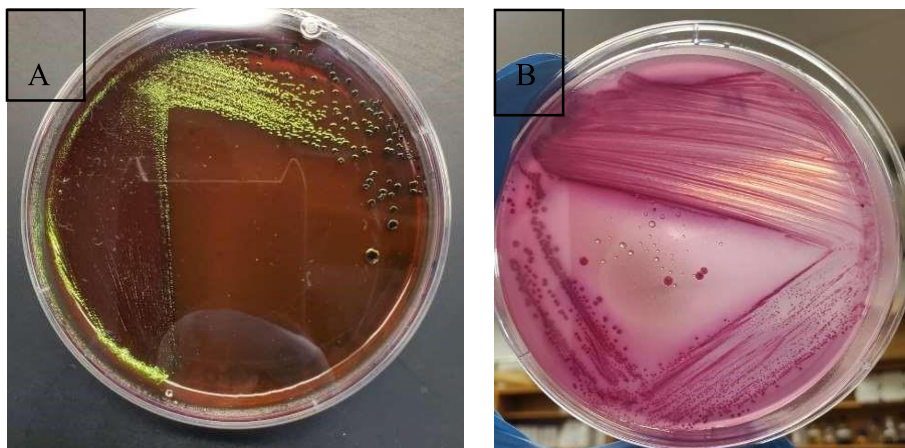


Figure 12. Presence of *E. coli* on A. EMB plate, B. MacConkey plate.



Figure 13. Restreaked isolated colonies on EMB

***E. coli* Identification by MALDI-TOF MS**

Isolates tested positive for *E. coli* on EMB, and MacConkey were restreaked on TSA plates for MALDI-TOF MS. Isolates from sewage, septic tank, seal scat, and dog feces were spotted in triplicates onto a single 384 spot steel target. Upon receiving the spectra, MTU table (Appendix) was obtained by the R-script, differentiating isolates with different MTU's. From 100 isolates, 70 were found to be *E. coli* with a score value of more than 2.0, the other 30 isolates included *Enterococci*, *Klebsiella pneumonia*, and *Pseudomonas* spp (Fig. 14). The dendrogram shows a distinction between libraries generated from sewage and septic tank from libraries generated from animal sources.

MALDI-TOF MS did not clearly distinguish libraries generated from seal and canine scat.

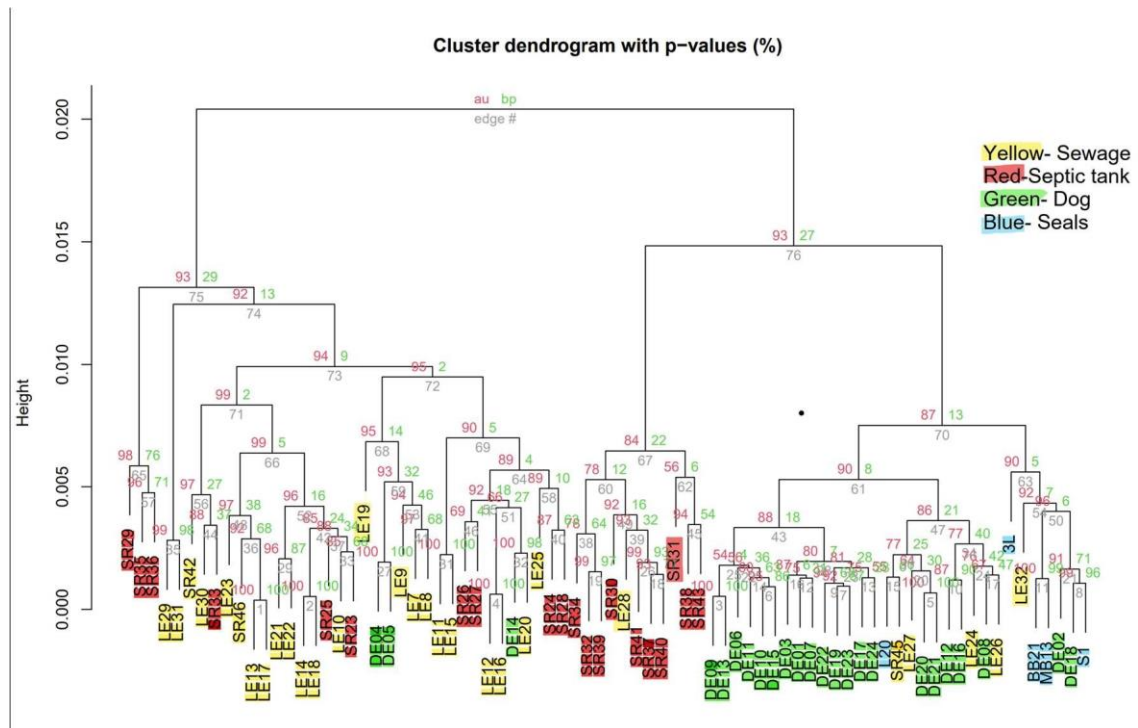


Figure 14. Dendrogram for *E. coli* from multiple sources; Sewage, Septic tanks, Dog feces and Seal scat, with bootstrap values in red numbers.

DNA Extraction

Isolate DNA extraction

Inspection of agarose gels of DNA extracted from pure cultures indicated the presence of either RNA or small fragment DNA, due to presence of smaller size bands at the bottom of the gel (Fig. 15,16). Nanodrop readings also suggested contamination, the A_{260}/A_{280} ratios were above 1.8, which is consistent with RNA contamination.

Table 7. Nanodrop readings for DNA extraction by for each isolate

Sample ID	DNA concentration ng/μl	Volume(μl)	Yield (ug)	A260/280	A260/230
3L	1698	35	59	2.09	1.95
MB13	1348	35	47	2.05	1.81
ST21	2186	35	76	2.08	1.86
ST28	1334	35	46	1.97	1.47
TA04	119	35	6	1.97	1.45
BB21	2051	35	71	2.11	2.02
L20	1462	35	51	2.05	1.84
TA01	1355	35	47	1.99	1.48
LE24	1219	35	42	2.04	1.82
LE27	1301	35	45	2.01	1.73
LE30	1203	35	42	1.92	1.32
LE34	1197	35	41	2.02	1.54
SR40	1615	35	56	2.03	1.79
SR44	1127	35	39	2.03	1.30
SR48	77	35	2	1.72	0.70
SR49	112	35	3	1.76	1.45
SR46	70	35	2	1.72	0.74
SR31	1103	35	38	2.03	1.65
SR34	451	35	15	1.99	1.37
SR27	2056	35	71	2.10	2.14
DE06	1046	35	36	2.02	1.72
DE04	1167	35	40	1.91	1.41
DE20	1050	35	36	2.08	1.91
DE17	1064	35	37	1.98	1.65

Table 8. Samples after DNA extraction added in each lane of agarose gel with 5 μ l volume in each well.

Lane	Sample
L-Ladder	1KB DNA Ladder
GEL 1	
1	LE24
2	LE30
3	BB21
4	MB13
5	L20
6	TA04
L-Ladder	1KB DNA Ladder
GEL 2	
L-Ladder	1KB DNA Ladder
1	SR27
2	SR40
3	SR31
4	SR34
5	SR21
6	SR28
7	SR46
8	SR48
L-Ladder	1KB DNA Ladder
L-Ladder	1KB DNA Ladder

9	SR44
10	SR49
11	DE20
12	DE17
13	DE06
14	DE04
15	LE27
16	LE34
Ladder	1KB DNA Ladder

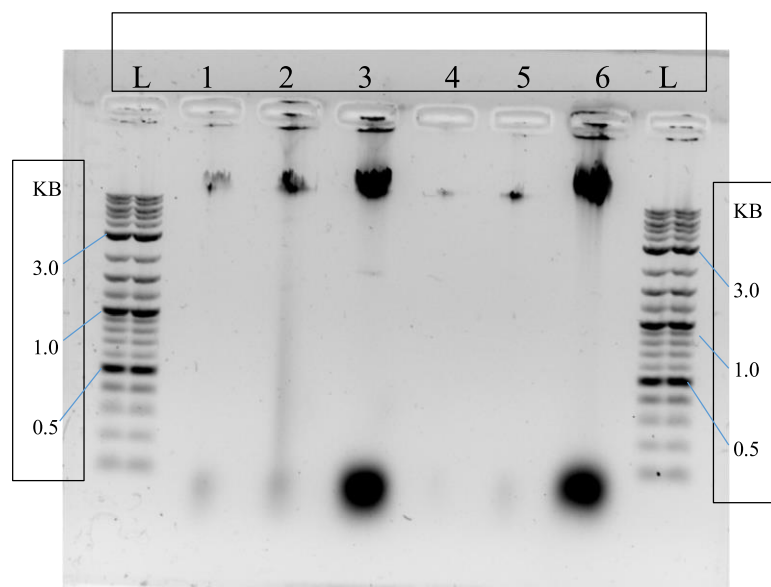


Figure 15. Gel 1, DNA from isolates (1% Agarose gel, in 1X TBE, Gel Red, 150V, 45 mins, 1kb plus DNA ladder). Showing high molecular bands above 10 kb with some smaller bands at the bottom of the gel indicating small fragment and RNA contamination.

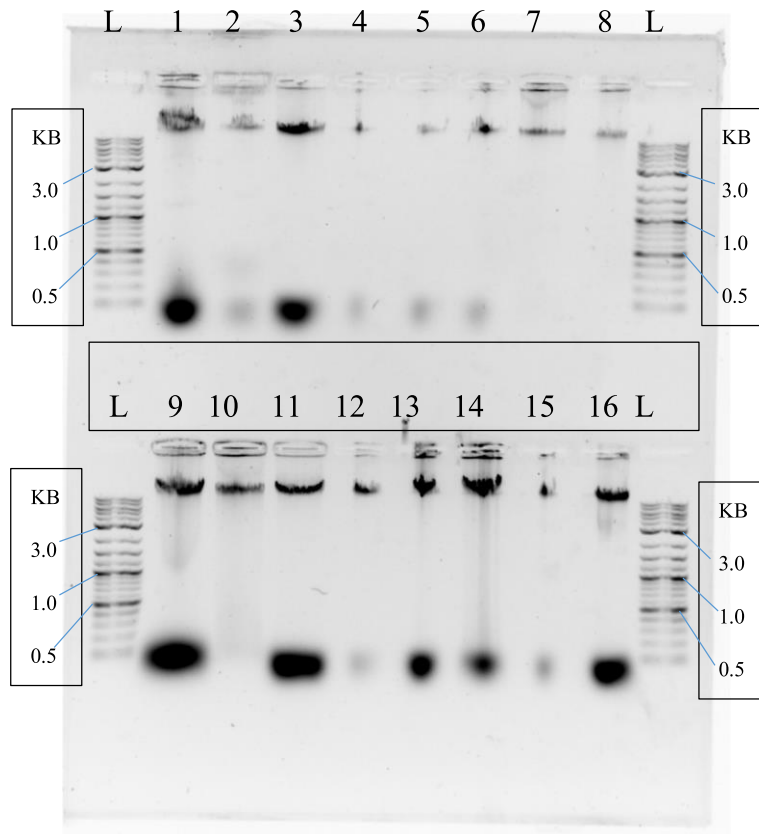


Figure 16. Gel 2, DNA from isolates (1% Agarose gel, in 1X TBE, Gel Red, 150V, 45 mins, 1kb plus DNA ladder). Showing high molecular bands above 10 KB with some smaller bands at the bottom of the gel indicating small fragment and RNA contamination

DNA Concentration

Purification to remove small fragments or RNA contamination appeared effective. Agarose gels after purification with magnetic beads showed little to no contamination with small fragments (Fig. 17). These agarose gels showed high molecular DNA suitable for genomic analysis.

Table 9. Nanodrop readings after DNA purification with magnetic beads

Isolate	A260/280	A260/230	conc ng/ul	Yield (ng)
3L	1.86	1.68	72.2	2527
L20	1.64	0.83	26.5	928
MB13	1.71	1.01	176.7	6184
BB21	1.79	1.15	56.2	1967
DE20	1.76	1.18	65.2	2282
DE04	1.69	0.92	127.2	4452
SR27	2.00	2.08	216.4	7574
LE24	1.58	0.74	43.2	1512
LE30	1.84	1.75	58.8	2058
SR31	1.7	0.88	122.3	4281
ST28	1.67	0.86	225.6	7896
TA01	1.64	0.74	110.9	3882
ST21	1.75	1.06	213.3	7466
SR44	1.98	1.53	131.5	4603
SR49	1.84	1.31	19.9	697

Table 10. Samples after DNA concentration added in each lane of agarose gel with 5 μ l added to each well.

Lane	Sample
L-Ladder	1KB DNA Ladder
1	ST28
2	SR21
3	DE20
4	DE04
5	SR31
6	SR49
7	SR44
8	SR27
L-Ladder	1KB DNA Ladder
L-Ladder	1KB DNA Ladder
9	LE24
10	LE30
11	MB13
12	BB21
13	L20
14	3L
15	TA01
16	TA04
Ladder	1KB DNA Ladder

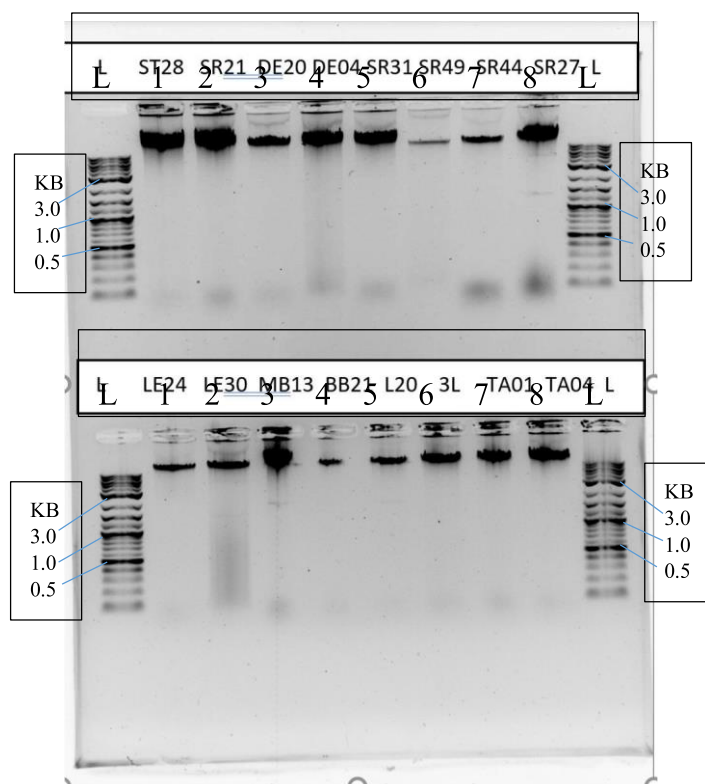


Figure 17. DNA from isolates after concentration by magnetic beads (1% Agarose gel, in 1X TBE, Gel Red, 150V, 45 mins with 1kb plus DNA ladder). Gel showed much lesser contamination by small fragment DNA and RNA.

Total DNA Extraction

Total (metagenomic) DNA was extracted from samples processed with DESS. DNA extracted from sewage and OWTS gave relatively pure nucleic acids (Table 11). The purity of A_{260}/A_{280} of the DNA extracted was near 1.8, indicating DNA extracted was free of RNA contamination. A_{260}/A_{230} was near 1.8 for seal tube number #2 and sewage, indicating that they are free of contamination. UV spectra generated from other samples were above 1.8, which suggests the presence of carbohydrates or glycogen residues. Based on agarose gel results, high molecular DNA obtained in all the samples was contaminated with RNA and/or small fragment DNA. (Fig. 18).

Table 11. Meta samples extraction added in each lane of agarose gel with 5 μ l of volume in each well.

Lane	Sample
L-Ladder	1 Kb DNA ladder
1	Sewage
2	Septic tank
3	Seal 1
4	Seal 2
5	Seal 3
6	Dog
L-Ladder	1 Kb DNA Ladder

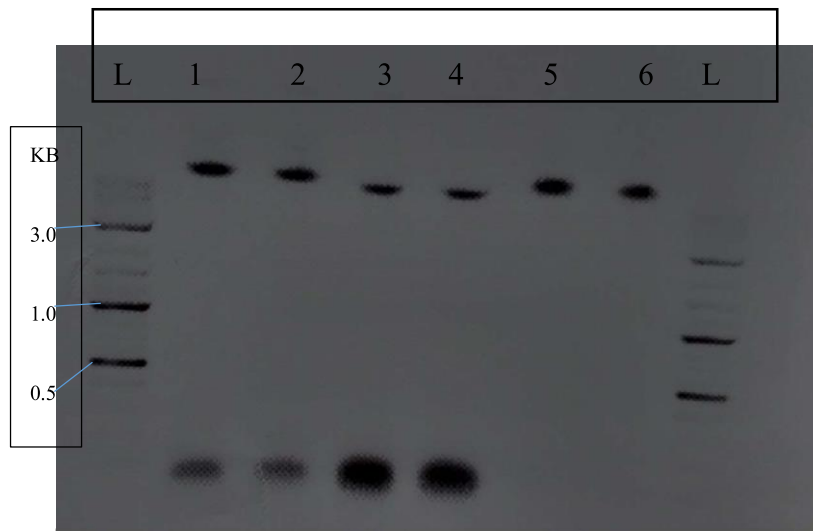


Figure 18. Total DNA extraction from meta-samples (1% agarose gel made in 1X TBE, Gel Red, 150V, 45 mins, 1Kb DNA Ladder). Gel showed high molecular bands above 10 kb for all samples.

Table 12. Nanodrop readings for total DNA extracts from meta-samples

Sample ID	DNA ng/μl	Volume (μl)	Yield (ug)	A260/280	A260/230
Seal #1	31	50	1530	1.97	0.21
Seal #2	213	50	10635	1.89	1.54
Seal #3	37	50	1840	1.75	0.21
Sewage WWTP	246	50	12295	1.89	1.54
Septic tank	112	50	5600	1.92	0.94
Dog feces	74	50	3680	1.89	0.87

DNA Quantification

AccuBlue standard graph showed a good fit to a linear model, which was used to calculate unknown sample concentrations (Table 13). Concentration for waste and fecal samples as assessed by from AccuBlue® agreed with concentration assessed by Nanodrop spectrometry (Fig. 19). Samples that showed high molecular DNA (Table 13) and were chosen for antibiotic resistance tests.

Table 13. DNA quantification by AccuBlue®, shows fluorescence value by 96-well plate reader along with the samples nanodrop readings. Yellow highlights were the isolates used for antibiotic resistance assays.

Isolate ID	AccuBlue ng/ul	A260/280	A260/230	Nanodrop ng/ul	Yield (ng)
3L	50	1.86	1.68	72	2527
BB21	21	1.79	1.15	56	1967
DE04	66	1.69	0.92	127	4452
DE20	36	1.76	1.18	65	2282
Dog	81	1.89	0.87	74	3680
L20	-2	1.64	0.83	27	928
LE24	18	1.58	0.74	43	1512
LE30	43	1.84	1.75	59	2058
MB13	171	1.71	1.01	177	6185
Seal 1	-9	1.97	0.21	31	1530
Seal 2	115	1.89	1.54	212	10635
Seal 3	25	1.75	0.21	37	1840
Septic	10	1.92	0.94	112	5600
Sewage 1	193	1.89	1.54	246	12295
ST21	157	1.75	1.06	213	7465
SR27	144	2.00	2.08	216	7574
SR31	89	1.7	0.88	122	4281
SR44	10	1.98	1.53	132	4603
SR49	2	1.84	1.31	20	697
ST28	179	1.67	0.86	226	7896
TA01	89	1.64	0.74	111	3882

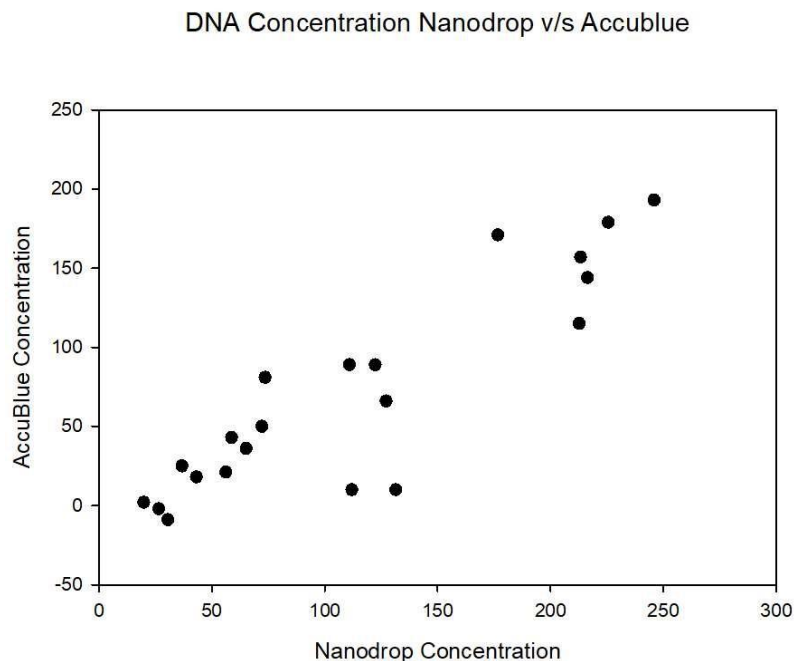


Figure 19. Comparison between concentration of DNA from Nanodrop (x-axis) and concentration of DNA from AccuBlue® (y-axis) made in SigmaPlot 14.5

Antibiotic Resistance

Kirby Bauer for Representative Microorganism

Each antibiotic has its own zone of resistance, susceptibility, and an intermediate range for different organisms (Hudzieki 2009). First antibiotics were tested on different species isolates from sewage (Appendix), which included both Gram- positive and Gramnegative bacteria. All isolates were found to be resistant to ampicillin, and erythromycin. Isolates were least resistant to chloramphenicol and tetracycline.

E. coli Antibiotic Resistance Broad Spectrum

Antibiotics used for E. coli isolates were chloramphenicol, ampicillin, tetracycline, erythromycin, streptomycin, and gentamycin. Each antibiotic has its acceptance zone of clearance, which defines its susceptibility towards bacteria, as observed in Table 16 (Hudzieki 2009). All E. coli isolates were found to be resistant to antibiotics streptomycin, erythromycin, and ampicillin. Least resistance was observed for tetracycline (Fig. 20)

E. coli Antibiotic Resistance Narrow Spectrum

Polymyxin B resistance was conducted using the Kirby- Bauer protocol. The zone of susceptibility (the distance between the zone of clearance) for E. coli was above 11 mm (Bijayini et al. 2010). A strain (SR27) from the septic tank sample showed resistance towards this antibiotic (Fig. 22).

Table 14. Antibiotic resistance data for E. coli from multiple sources with MTU and the zone of inhibition

Organism	Isolate	MTU	Antibiotic	Zone (mm)	Interpretation
E. coli	ST28	1	Chloramphenicol	20	Susceptible
			Ampicillin	5	Resistant
			Tetracycline	19	Susceptible
			Gentamicin	20	Susceptible
			Erythromycin	10	Resistant
			Streptomycin	14	Resistant
E. coli	DE04	3	Chloramphenicol	0	Resistant
			Ampicillin	0	Resistant
			Tetracycline	9	Resistant
			Gentamicin	0	Resistant
			Erythromycin	0	Resistant
			Streptomycin	0	Resistant
E. coli	SR49	8	Chloramphenicol	18	Susceptible
			Ampicillin	0	Resistant
			Tetracycline	21	Susceptible
			Gentamicin	0	Resistant
			Erythromycin	0	Resistant
			Streptomycin	0	Resistant
E. coli	LE30	6	Chloramphenicol	0	Resistant
			Ampicillin	0	Resistant
			Tetracycline	19	Susceptible

			Gentamicin	16	Susceptible
			Erythromycin	0	Resistant
Organism	Isolate	MTU	Antibiotic	Zone (mm)	Interpretation
			Streptomycin	0	Resistant
E. coli	SR27	4	Chloramphenicol	0	Resistant
			Ampicillin	0	Resistant
			Tetracycline	0	Resistant
			Gentamicin	20	Susceptible
			Erythromycin	0	Resistant
			Streptomycin	14	Resistant
E. coli	ST21	1	Chloramphenicol	15	Susceptible
			Ampicillin	0	Resistant
			Tetracycline	18	Susceptible
			Gentamicin	17	Susceptible
			Erythromycin	0	Resistant
			Streptomycin	10	Resistant
E. coli	MB13	2	Chloramphenicol	19	Susceptible
			Ampicillin	0	Resistant
			Tetracycline	21	Susceptible
			Gentamicin	16	Susceptible
			Erythromycin	8	Resistant
			Streptomycin	12	Resistant

Table 14 cont. Kirby Bauer protocol for Polymyxin B resistance on *E. coli* isolates and for direct samples from seal and sewage.

Organism	Isolate ID	MTU	Antibiotic	Zone (mm)	Interpretation
<i>E. coli</i>	SR27	4	Polymyxin B	0	Resistant
<i>E. coli</i>	LE34	6	Polymyxin B	20	Susceptible
<i>E. coli</i>	MB13	2	Polymyxin B	15	Susceptible
<i>E. coli</i>	ST21	1	Polymyxin B	17	Susceptible
Seal	Meta		Polymyxin B	24	Susceptible
Sewage	Meta		Polymyxin B	0	Susceptible

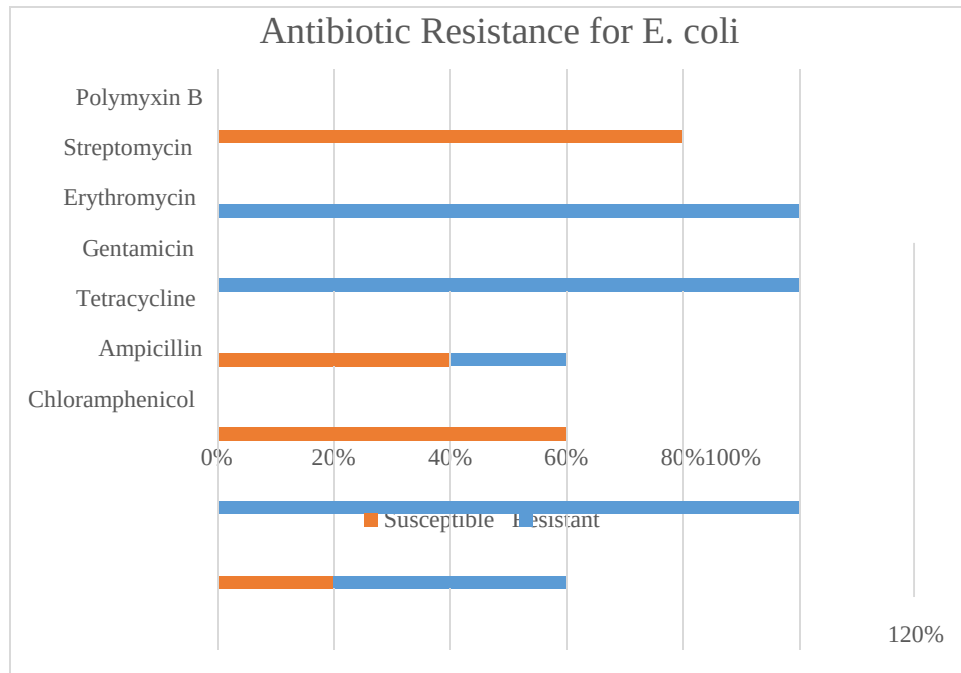


Figure 20. Resistance for *E. coli* in conjunction with different antibiotics.

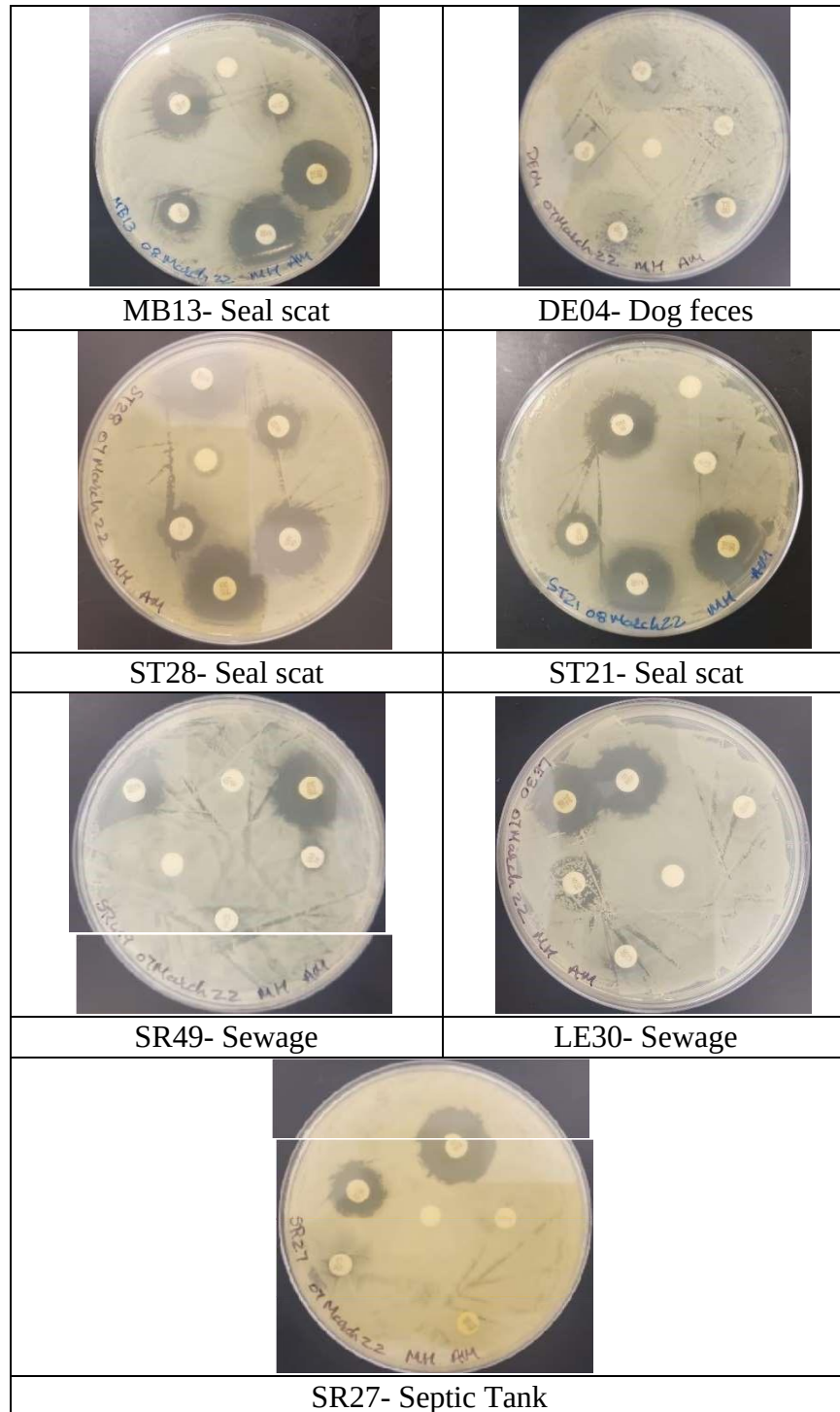


Figure 21. Antibiotic resistance plates with 6 different antibiotic discs in each plate showing zone of clearance for susceptibility.

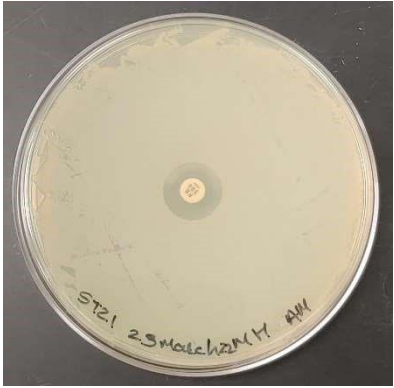
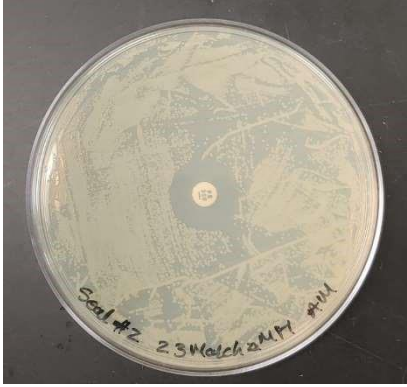
	
SR27- Septic Tank E. coli	LE34- Sewage E. coli
	
MB13- Seal E. coli	ST21- Seal E. coli
	
Meta-sample Seal	Meta-sample Sewage

Figure 22. Polymyxin B Antibiotic Resistance by Kirby-Bauer on *E. coli* and for direct samples from seal and sewage.

Colistin Resistance Evaluation by PCR

PCR was done using 16S rRNA primers, mcr-1, mcr-2, and mcr-3 primers and a positive control with G-Block. G-Block is a concatenation of predicted amplicons with primers for mcr-1, mcr-2, mcr-3. First, gel electrophoresis was conducted for products generated with 16S rRNA primers, which showed robust amplification for all the samples (Fig. 23). Mass of bands observed were as expected, as the 16S rRNA primer product length is 1.5 KB and samples also resulted in an amplicon of 1.5 KB. This indicates DNA preparations were PCR-amplifiable. For mcr-1 primers, there was positive thick band for G-Block near 0.3 kb, which was expected. Sewage meta-sample sample and SR27 (OWTS sample) showed larger size molecular bands than expected for mcr-1 primers. (Fig. 24). There were bands observed for other samples as well, but these were light and could be due to DNA shearing. For mcr-2 primers, gel electrophoresis showed bands of the predicted size only for the G-Block control. All bands generated from samples with mcr-2 primers were less than 0.2 KB (Fig. 25). This is most likely due to unincorporated primers.

However, mcr-3 primers did yield band of the predicted size from sewage sample (Fig. 26).

Table 15. Samples after DNA amplification with 16S primer added in each lane of agarose gel with 5 μ l added to each well.

Lane	Sample
L	1 KB DNA Ladder
1	G-Block
2	ST21
3	LE34
4	SR27
5	MB13
6	Sewage sample
7	Seal scat

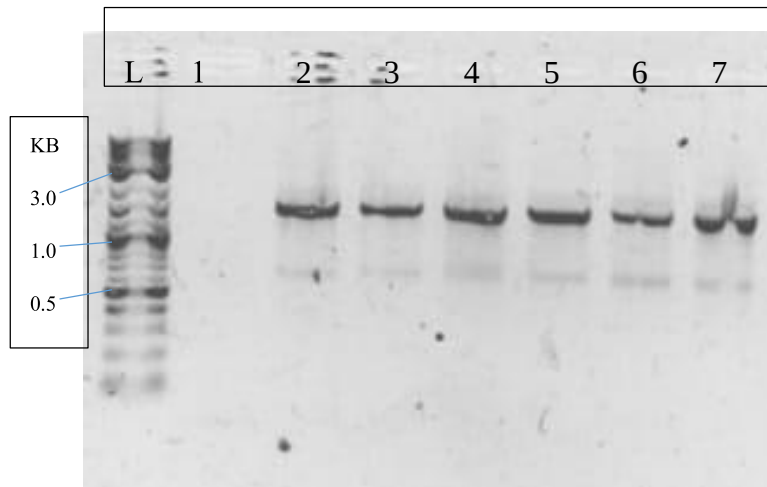


Figure 23. PCR Amplification with 16S rRNA gene primers (1% Agarose gel with 1X TBE, Gel Red, 150V, 45 mins with 1 kb DNA plus Ladder), shows amplification for all isolates.

Table 16. Samples after DNA amplification with *mcr-1* primer added in each lane of agarose gel with 5 μ l added to each well.

Lane	Sample
L	1 KB DNA Ladder
1	G-Block
2	ST21
3	MB13
4	SR27
5	LE34
6	Sewage sample
7	Seal scat

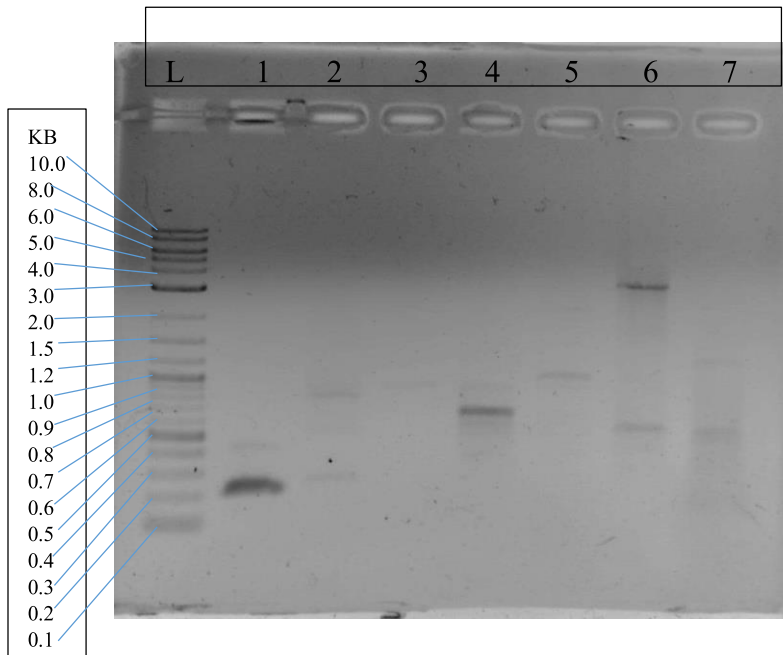


Figure 24. PCR Amplification with *mcr-1* primers (1% Agarose gel with 1X TBE, Gel Red, 150V, 45 mins with 1 kb DNA plus Ladder), shows unexpected bands for sewage (lane 6) and SR27 (lane 4) samples, when compared with positive control (lane 1).

Table 17. Samples after DNA amplification with *mcr-2* primer added in each lane of agarose gel with 5 μ l added to each well.

Lane	Sample
L	1 KB DNA Ladder
1	G-Block
2	ST21
3	LE34
4	SR27
5	MB13
6	Sewage sample
7	Seal scat

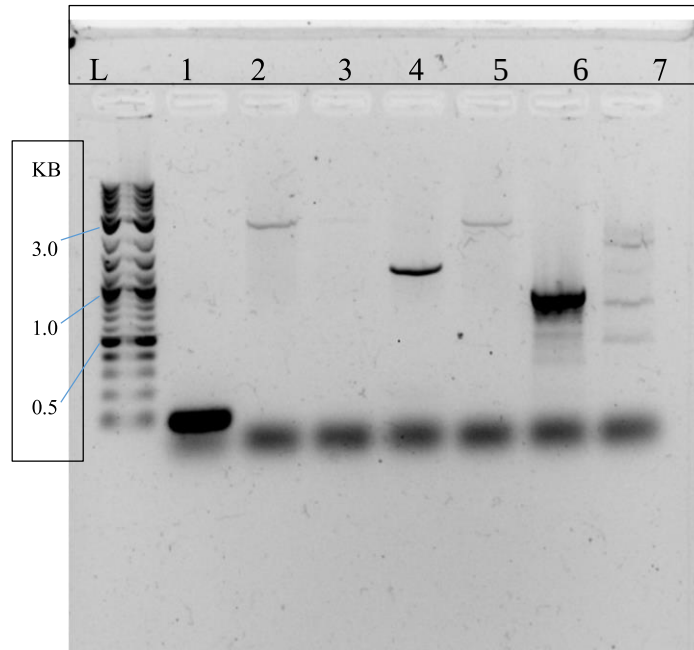


Figure 25. PCR Amplification with *mcr-2* primers (1% Agarose gel with 1X TBE, Gel Red, 150V, 45 mins with 1 kb DNA plus Ladder), gel shows expected results with low molecular bands for all samples. Unexpected bands for sewage (lane 6) and SR27 (lane 4) samples, when compared with G-block (lane 1).

Table 18. Samples after DNA amplification with *mcr-3* primer added in each lane of agarose gel with 5 μ l added to each well.

Lane	Sample
L	1 KB DNA Ladder
1	G-Block
2	ST21
3	LE34
4	SR27
5	MB13
6	Sewage sample
7	Seal scat

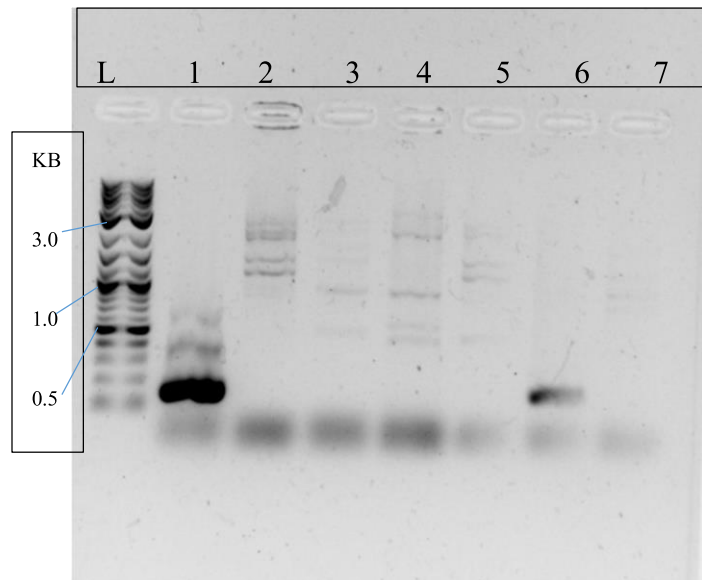


Figure 26. PCR Amplification with *mcr-3* primers (1% Agarose gel with 1X TBE, Gel Red, 150V, 45 mins with 1 kb DNA plus Ladder), gel shows expected results for sewage sample (lane 6) as compared to G-block (lane 1).

CHAPTER IV:

DISCUSSION

The study was conducted to evaluate MALDI-TOF MS as a method for multiple source tracking. *E. coli* strains were cultured from raw sewage and OWTSS samples and from animal sources using selective and deferential primary and secondary media. These *E. coli* isolates were sorted into human and animal specific cluster using MALDI-TOF MS, which suggests this mass spectrometry method is suitable for bacterial source tracking of *E. coli* from sewage and OWTSS. However, MALDI-TOF MS could not clearly differentiate between animal sources of *E. coli*.

Antibiotic resistance of representative isolated was assessed by Kirby- Bauer and PCR. All tested *E. coli* isolated were all resistant to ampicillin, streptomycin, and gentamicin. Only an *E. coli* isolated from OWTSS was resistant to polymyxin B. This is consistent with the observation that only the sewage meta-sample tested positive for the presence of a gene with colistin resistance.

E. coli Isolation

In preliminary work, isolation directly to differential and selective media (EMB and MacConkey) did not consistently yield *E. coli* isolates. Faster growing species appeared to take over the plates. Even with the dilution of 1:100 with 1X PBS, there were only 2 out of 103 isolates identified as *E. coli*. Therefore, A1 media was used as the primary media to isolate *E. coli*. This differential media and selective media is not widely used or EPA recommended. Most water quality managers use the Colitert kit (IDEXX, Westbrook, ME) for quantifying *E. coli* in water samples. A1 and the Colitert kit may select for different *E. coli* strains.

MALDI-TOF MS

The primary aim for the project was to determine if MALDI-TOF can be exploited for tracking FIB from sewage and OWTSS. MALDI-TOF appears suitable for

bacterial source tracking of *E. coli* from sewage and OWTs sources. However, did not differentiate *E. coli* strains isolated from seal scat and dog waste. These results agree with an earlier study that showed MALDI-TOF can distinguish *E. coli* strains isolated from sewage from strains isolated from bird droppings (Siegrist et al. 2007). The small sample size of this study limits the interpretation of the results. Wastewater, both sewage and OWTs represent composites of hundreds of thousands or dozens of people and there is little variation between *E. coli* isolated from sewage worldwide (Stoppe et al. 2017). However, only two dogs and one group of captive seals were sampled herein.

In the cluster dendrogram by MALDI-TOF MS in this study, a difference between strains isolated from sewage and OWTs was observed, although both systems contain human waste. This could reflect how the systems function. Sewage treatment plants process human waste within a day or two and the waste is vigorously agitated and aerated. OWTs are not agitated and typically accumulate waste for months or years. This long period of storage gives microbes in the system ample opportunity to exchange ARG.

MALDI-TOF MS has advantages over WGS for microbial source tracking. WGS requires skilled personnel and is relatively expensive. MALDI-TOF is cheap and easy. 16S rRNA gene sequencing is also inexpensive and easy, and widely used for bacterial identification; however, this approach lacks the resolution of MALDI-TOF MS and WGS. For example, in a study on 16S gene sequences generated from *E. coli* strains isolated from humans and cows, 16S rRNA gene sequencing was unable to differentiate *E. coli* isolates by source. Indeed some the strains of *E. coli* showed 100% similarity to *S. dysenteriae* (Suardana 2014).

The inability of MALDI-TOF MS to distinguish animal sources is consistent with phylogenomic evidence. Sewage-sourced *E. coli* strains fall into four phylogroups (A, B1, B2, and D) that can be differentiated with a multiplexed PCR reaction (Clermont et al. 2012). Groups A and B2 are found in waterways contaminated with human waste, B1

with domestic animals, and D with pristine sites (Stoppe et al. 2014). To evaluate this approach, I analyzed WGS sequences generated from *E. coli* strains from human, companion animal and aquatic animals with functions in the EnteroBase pipeline. Similar results were obtained as with MALDI-TOF MS. Human waste was differentiated from animal sources, as clade A was only associated with human waste, but there was no differentiation observed between animal sources (Fig. 27).

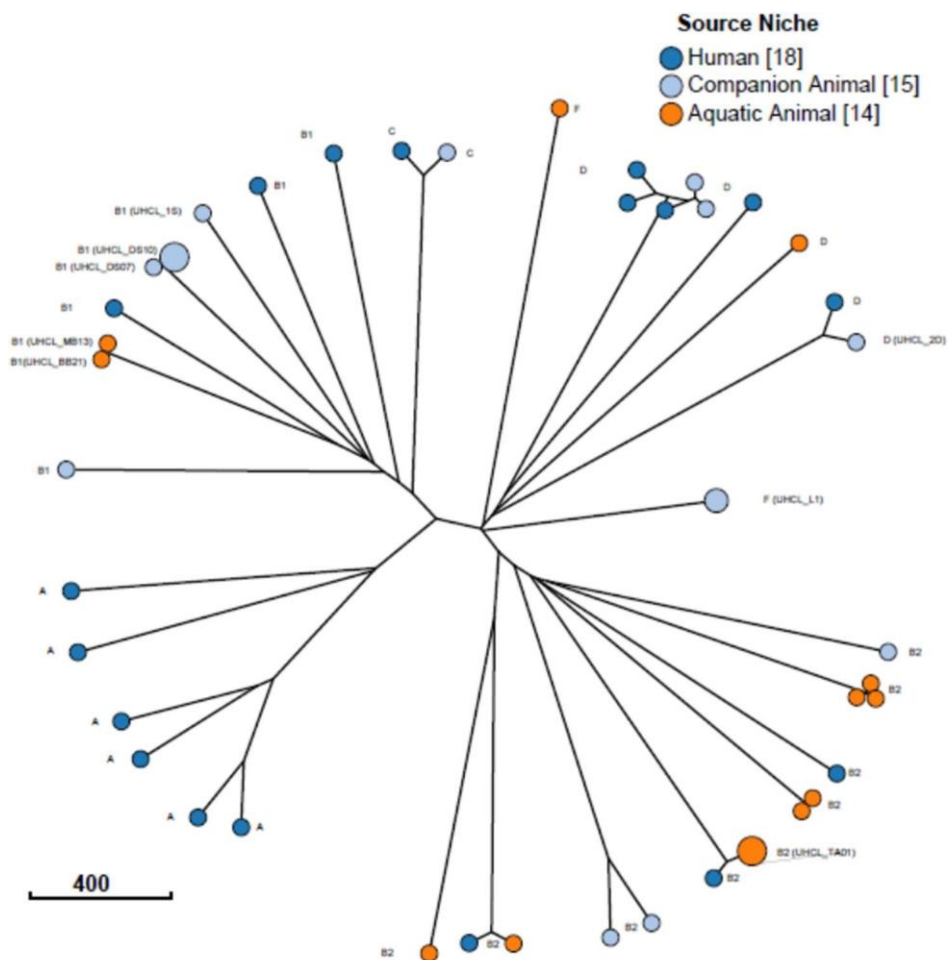


Figure 27: Hierarchical clustering based on core genome multilocus sequence typing. Clusters generated from *E. coli* strains isolated from humans, companion animal and aquatic animal. Figure was generated with GrapeTree (Zhou, 2018) in Enterobase (Zhou, 2020)

Antibiotic Resistance

Bacteria isolated from multiple sources all showed resistance to a range of broad spectrum and narrow spectrum antibiotics. Resistance to polymyxin B was only observed for *E. coli* isolated from OWTS and only the meta-sample DNA generated from sewage tested positive for the presence of genes associated with resistance to colistin, as assessed with mcr-3 primers. This suggests that bacteria that carry the polymyxin resistance genes are prevalent in human waste, and these ARB could be indicative of human waste.

Resistance to ampicillin, streptomycin and erythromycin is common in *E. coli* (Tadesse et al. 2012). Resistance to polymyxin B is relatively rare. For example, 1% of *E. coli* strains isolated from various sources in China were resistant to polymyxin (Yan et al. 2021). Only mcr-3 primers were effective in generating the expected PCR product. Primers for mcr-1 yielded larger than expected size bands for sewage and septic tank samples, and other samples had fainter bands for other isolates. The results for mcr-1 and mcr-2 primers were inconclusive, which could reflect contaminants in sewage samples that lead to spurious amplification products (Lorenz 2012). Smaller bands were observed in comparison with positive control (G-Block) for other isolates for mcr-2 which could be due to addition of excessive primers in the reaction.

Limitations

A much broader Bruker reference database is required of MALDI-TOF to reliably identify all isolates, as well as to differentiate between animal sources. This can be improved in time by addition of more data to Bruker diagnostics or a suitable database. DNA quantification technique could be improved by using AccuBlue® Green with Qubit 4, as AccuBlue® is time consuming and has multiple steps in making standards from stock following the protocol in 96 well plates which leads to errors. AccuBlue® could not be performed in triplicates due to shortage of materials and supply chain issues. Quantification of the DNA for antibiotic genes could not be performed by qPCR as the

master mix for the same was not received with in time. However, in the future qPCR could be used to quantify the antibiotic genes present in the isolate.

CHAPTER V:

CONCLUSION AND FUTURE DIRECTION

MALDI-TOF MS can differentiate human and animal sources of *E. coli* quickly and inexpensively. Isolates from other animal sources, including cats, cattle, feral hogs, and deer, should also be analyzed. These animals are common in the Houston-Galveston area and therefore can be a source of fecal contamination.

The Houston-Galveston area is also home to a number of national medical centers. Isolates from hospital waste would play an important role in knowing the emergence of resistance coming into sewage (Janda et al. 2007, Jenkins et al. 2012, Wang et al. 2015). Extreme weather events could spread ARB, including virulent *E. coli* strains, from wastewater treatment plants that receive waste from these medical centers.

MALDI-TOF MS enabled microbial source tracking could complement widely used qPCR approaches to ARG, including protocols for *sul1*, *sul2*, *tetG*, *tetO*, *tetW*, *bla*. (Devarajan et al. 2016, Garner et al. 2016, Yun-Wen et al. 2015, Gorecki et al. 2019). This study suggests qPCR protocols for *mcr-3* should be developed too. WGS could also be used to understand similarity between isolates characterized in this study with the published literature (Fresia et al. 2019, Vos 2017).

REFERENCES

- Aridi R, Tuchscherer R, & Malakoff, D (2020) Tales from the storm: how four scientists tracked Hurricane Laura. Science Insider.
<https://www.insider.com/hurricanelaura-climate-change-warming-water-2020-8>
- Barreto A, Guimarães B, Radhouani H, Araújo C, Gonçalves A, Gaspar E, Rodrigues J, Igrejas G, & Poeta P (2009) Detection of antibiotic resistant *E. coli* and *Enterococcus* spp. in stool of healthy growing children in Portugal. *Journal of Basic Microbiology* 49:503–512.
- Behera B, Mathur P, Das A, Kapil A, Gupta B, Bhoi S, Farooque K, Sharma V, Misra M.C (2009) Evaluation of susceptibility testing methods for polymyxin. *International Journal of Infectious Diseases* 7:596-601.
doi:10.1016/j.ijid.2009.09.001
- Bian X, Liu X, Hu F, Feng M, Chen Y, Bergen P. J, Li J, Li X, Guo Y, & Zhang J. (2022). Pharmacokinetic/Pharmacodynamic based breakpoints of polymyxin B for bloodstream infections caused by multidrug-resistant Gram-negative pathogens. *Frontiers in Pharmacology* 12:785893.
- Brown B, Allard M, Bazaco MC, Blankenship J, Minor T (2021) An economic evaluation of the whole genome sequencing source tracking program in the U.S. *PLOS ONE* 16:e0258262
- de Castro Stoppe N, Silva JS, Torres TT, Carlos C, Hachich EM, Sato MI, Saraiva, AM, & Ottoboni LM (2014) Clustering of water bodies in unpolluted and polluted environments based on *Escherichia coli* phylogroup abundance using a simple interaction database. *Genetics and Molecular Biology* 37:694–701.
- CDC, Centers for Disease Control and Prevention (2021), Antibiotic / Antimicrobial Resistance (AR / AMR). <https://www.cdc.gov/drugresistance/index.html>

- CDC, Centers for Disease Control and Prevention (2019), Antibiotic resistance threats in the United States. <https://www.cdc.gov/drugresistance/pdf/threats-report/2019-arthreats-report-508.pdf>
- Chalmers G, Davis KE, Poljak Z, Friendship R, Mulvey MR, Deckert AE, Reid-Smith RJ, & Boerlin P (2018) A method to detect *Escherichia coli* carrying the colistinresistance genes *mcr-1* and *mcr-2* using a single real-time polymerase chain reaction and its application to chicken cecal and porcine fecal samples. *Canadian Journal of Veterinary Research* 82:312–315.
- CHRON NEWS, Stuckey A (2017) Raw sewage spilled in Houston area after wastewater plants damaged by Harvey. <https://www.chron.com/news/houston-texas/article/Nearly-a-dozen-wastewater-treatment-facilities-12209605.php>
- Clermont O, Christenson JK, Denamur E, & Gordon DM (2013) The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environmental Microbiology Reports* 5:58–65.
- Converse RR, Blackwood AD, Kirs M, Griffith JF, Noble RT (2009) Rapid QPCR-based assay for fecal *Bacteroides* spp. as a tool for assessing fecal contamination in recreational waters. *Water Research* 43:4828-4837.
- D'Andrea MM, Fraziano M, Thaller MC, & Rossolini GM (2019) The urgent need for novel antimicrobial agents and strategies to fight antibiotic resistance. *Antibiotics* 8:254.
- Davies J (2006) Where have all the antibiotics gone? *The Canadian Journal of Infectious Diseases & Medical Microbiology* 17:287–290.
- Devarajan N, Laffite A, Mulaji CK, Otamonga JP, Mpiana PT, Mubedi JI, Prabakar K, Ibelings BW, Poté J (2016) Occurrence of antibiotic resistance genes and bacterial markers in a tropical river receiving hospital and urban wastewaters. *PloS ONE* 11: e0149211.

- EPA, United States Environment Protection Agency, Water: Monitoring & Assessment, 5.11 Fecal Bacteria. <https://archive.epa.gov/water/archive/web/html/vms511.html>
- Eramo A, Delos Reyes H, & Fahrenfeld NL (2017) Partitioning of antibiotic resistance genes and fecal indicators varies intra and inter-storm during combined sewer overflows. *Frontiers in Microbiology* 8:20-24.
- Farnaz M (2021) Application of whole genome sequencing and MALDI-TOF to identification of *Bacillus* species isolated from cleanrooms at NASA Johnson Space Center. Master Thesis, University of Houston- Clear Lake
- Field KG, Chern EC, Dick LK, Fuhrman J, Griffith J, Holden PA, LaMontagne MG, Olson B, Simonich MT (2003) A comparative study of culture-independent, library-independent genotypic methods of fecal source tracking. *Journal of Water Health* 1:181-194.
- Fresia P, Antelo V, Salazar C, Giménez M, D'Alessandro B, Afshinnkoo E, Mason C, Gonnet GH, & Iraola G (2019) Urban metagenomics uncover antibiotic resistance reservoirs in coastal beach and sewage waters. *Microbiome* 7:35.
- Garner E, Benitez R, Wagoner E, Sawyer R, Schaberg E, Hession WC, Krometis LH, Badgley BD, & Pruden A (2017) Stormwater loadings of antibiotic resistance genes in an urban stream. *Water Research* 123:144–152.
- Garner E, Wallace JS, Argoty GA, Wilkinson C, Fahrenfeld N, Heath LS, Zhang L, Arabi M, Aga DS, & Pruden A (2016) Metagenomic profiling of historic Colorado Front Range flood impact on distribution of riverine antibiotic resistance genes. *Scientific Reports* 6:38432.
- Gogry F, Siddiqui M, Sultan I, & Haq QM (2021) Current update on intrinsic and acquired colistin resistance mechanisms in bacteria. *Frontiers in Medicine* 8:677720

- Gomi R, Matsuda T, Matsui Y, Yoneda M (2014) Fecal source tracking in water by nextgeneration sequencing technologies using host-specific *Escherichia coli* genetic markers. *Environmental Science & Technology* 48:9616-9623.
- Gong B, Cao H, Peng C et al (2019) High-throughput sequencing and analysis of microbial communities in the mangrove swamps along the coast of Beibu Gulf in Guangxi, China. *Scientific Report* 9:9377
- Gorecki A, Decewicz P, Dziurzynski M, Janeczko A, Drewniak L, & Dziewit L (2019) Literature-based, manually curated database of PCR primers for the detection of antibiotic resistance genes in various environments. *Water Research* 161:211–221.
- Granado-Serrano AB, Martín-Garí M, Sánchez V, Riart Solans M, Berdún R, Ludwig IA, Rubió, L, Vilaprinyó E, Portero-Otín M, & Serrano J (2019) Faecal bacterial and short-chain fatty acids signature in hypercholesterolemia. *Scientific Reports* 9:1772.
- Hettiarachchi S, Wasko C, Sharma A (2018) Increase in flood risk resulting from climate change in a developed urban watershed – the role of storm temporal patterns. *Hydrology and Earth System Sciences* 22:2041-56.
- Hudzieki (2009) Kirby-Bauer disk diffusion susceptibility test protocol. American Society for Microbiology.
- Janda JM, & Abbott SL (2007) 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *Journal of Clinical Microbiology* 45:2761–2764.
- Jenkins C, Ling CL, Ciesielczuk HL, Lockwood J, Hopkins S, McHugh TD, Gillespie SH, & Kibbler CC (2012) Detection and identification of bacteria in clinical samples by 16S rRNA gene sequencing: comparison of two different approaches in clinical practice. *Journal of Medical Microbiology* 61:483–488.

- Kaufmann M, Lewandowski J, Choryński A, & Wiering M (2016) Shock events and flood risk management: a media analysis of the institutional long-term effects of flood events in the Netherlands and Poland. *Ecology and Society* 21:51.
- Kawecki S, Kuleck G, Dorsey JH et al (2017) The prevalence of antibiotic-resistant bacteria (ARB) in waters of the Lower Ballona Creek Watershed, Los Angeles County, California. *Environment Monitor Assess* 189:216.
- Knutson T, Camargo SJ, Chan JCL, Emanuel K, Ho C-H, Kossin J, Mohapatra M, Satoh M, Sugi M, Walsh K, Wu L (2020) Tropical cyclones and climate change assessment: Part ii: Projected response to anthropogenic warming. *Bulletin of the American Meteorological Society* 101:E303-E322.
- Li B, Yin F, Zhao X, Guo Y, Wang W, Wang P, Zhu H, Yin Y, & Wang X (2020) Colistin resistance gene *mcr-1* mediates cell permeability and resistance to hydrophobic antibiotics. *Frontiers in Microbiology* 10:3015.
- Li J, Shi X, Yin W, Wang Y, Shen Z, Ding S, & Wang S (2017) A multiplex SYBR green real-time PCR assay for the detection of three colistin resistance genes from cultured bacteria, feces, and environment samples. *Frontiers in Microbiology* 8:2078.
- Li Z, Cao Y, Yi L, Liu J. H, & Yang Q (2019) Emergent polymyxin resistance: end of an era? *Open Forum Infectious Diseases* 6:368.
- Lobanovska M, & Pilla G (2017) Penicillin's discovery and antibiotic resistance: lessons for the future? *The Yale Journal of Biology and Medicine* 90:135–145.
- Lorenz TC (2012) Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. *Journal of Visualized Experiments* 63:e3998
- Lucigen, MasterPure™ Complete DNA and RNA Purification Kit. Biosearch Technologies.

- LaMontagne MG, Tran PL, Benavidez A, Morano LD (2021) Development of an inexpensive matrix-assisted laser desorption—time of flight mass spectrometry method for the identification of endophytes and rhizobacteria cultured from the microbiome associated with maize. *PeerJ* 9:e11359.
- LaMontagne MG, Zhan Y, Guillen G, Gentry TJ, Allen MS (2022) Microbial community structure and fecal contamination of floodwaters generated by Hurricane Harvey. *bioRxiv*. 2022.01.20.476661.
- Lee KM, Adams M, Klassen JL (2019) Evaluation of DESS as a storage medium for microbial community analysis. *PeerJ* 7:e6414.
- MACHEREY-NAGEL, Genomic DNA from stool samples- NucleoSpin® DNA Stool 740472.10. 2016.
- Maier T, Klepel S, Renner U, & Kostrzewa M (2006) Fast and reliable Maldi-TOF MS—based microorganism identification. *Nature Methods* 3:1-2
- Moore DF, Harwood VJ, Ferguson DM, Lukasik J, Hannah P, Getrich M, Brownell M (2005) Evaluation of antibiotic resistance analysis and ribotyping for identification of faecal pollution sources in an urban watershed. *Journal of Applied Microbiology* 99:618-628.
- Motlagh AM, & Yang Z (2019) Detection and occurrence of indicator organisms and pathogens. *Water Environment Research* 91:1402-1408
- Nouws S, Bogaerts B, Verhaegen B, Denayer S, Laeremans L, Marchal K, Roosens NHC, Vanneste K, De Keersmaecker SCJ (2021) Whole genome sequencing provides an added value to the investigation of Staphylococcal food poisoning outbreaks. *Frontiers in Microbiology* 12:750278.
- Olaitan AO, Morand S, & Rolain JM (2014) Mechanisms of polymyxin resistance: acquired and intrinsic resistance in bacteria. *Frontiers in Microbiology* 5:643.

- Pandey PK, Kass PH, Soupir ML, Biswas S, & Singh VP (2014) Contamination of water resources by pathogenic bacteria. *AMB Express* 4:51.
- Paul DV, Thompson F, Thompson C, & Swings J (2017) Chapter 2: A flavor of prokaryotic taxonomy: systematics revisited, *Microbial Resources* 29-44.
- Pingfeng Y, Zaleski A, Li Q, He Y, Mapili K, Pruden A, Alvarez P, & Stadler L (2018) Elevated levels of pathogenic indicator bacteria and antibiotic resistance genes after hurricane harvey's flooding in Houston, *Environmental Science & Technology Letters* 5:481-486.
- Raza S, Kim J, Sadowsky MJ, Unno T (2021) Microbial source tracking using metagenomics and other new technologies. *Journal of Microbiology* 59:259-269.
- Reinthal FF, Posch J, Feierl G, Wüst G, Haas D, Ruckebauer G, Mascher F, & Marth E (2003) Antibiotic resistance of *E. coli* in sewage and sludge. *Water Research* 37:1685–1690.
- Schmitz S, & Suchodolski J (2016) Understanding the canine intestinal microbiota and its modification by pro-, pre- and synbiotics - what is the evidence? *Veterinary Medicine and Science* 2:71–94.
- Sericano J, Wade T, Brooks J, & Safe S (1994) Toxicological significance of non-, mono, and diortho-substituted polychlorinated biphenyls in oysters from Galveston. *Society of Environmental Toxicology and Chemistry* 13:1797-1803.
- Shah MS, Eppinger M, Ahmed S, Shah AA, Hameed A, & Hasan F (2016) Flooding adds pathogenic *Escherichia coli* strains to the water sources in southern Khyber Pakhtunkhwa, Pakistan. *Indian Journal of Medical Microbiology* 34:483–488.
- Siegrist TJ, Anderson PD, Huen WH, Kleinheinz GT, McDermott CM, & Sandrin TR (2007) Discrimination and characterization of environmental strains of *Escherichia coli* by matrix-assisted laser desorption/ionization time-of-flight mass

- spectrometry (MALDI-TOF-MS). *Journal of Microbiological Methods* 68:554–562.
- Silpe JE, Wong J, Owen SV, Baym M, & Balskus E. P (2022) The bacterial toxin colibactin triggers prophage induction. *Nature* 603:315–320.
- Singhal N, Kumar M, Kanaujia PK, & Viridi JS (2015) MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology* 6:791.
- Stoppe NC, Silva JS, Carlos C, Sato M, Saraiva AM, Ottoboni L, & Torres TT (2017) Worldwide Phylogenetic Group Patterns of *Escherichia coli* from Commensal Human and Wastewater Treatment Plant Isolates. *Frontiers in Microbiology* 8:2512.
- Suardana IW (2014) Analysis of nucleotide sequences of the 16S rRNA gene of novel *Escherichia coli* strains isolated from feces of human and Bali cattle. *Journal of Nucleic Acids* 2014:1-7.
- Tadesse DA, Zhao S, Tong E, Ayers S, Singh A, Bartholomew MJ, & McDermott PF (2012) Antimicrobial drug resistance in *Escherichia coli* from humans and food animals, United States, 1950-2002. *Emerging Infectious Diseases* 18:741–749.
- Ventola CL (2015) The antibiotic resistance crisis: part 1: causes and threats. *Journal for Formulary Management* 40: 277–283.
- Wang X, & Mayer LW (2015) A Phylogenetic perspective on molecular epidemiology, *Molecular Medical Microbiology* 517-536.
- Walters KE, & Martiny J (2020) Alpha-, beta-, and gamma-diversity of bacteria varies across habitats. *PloS ONE* 15:e0233872.
- Weitzman M (2018) Evaluation of omega Mag-Bind® totalpure NGS beads for DNA size selection. University of Oregon.

- https://gc3fstorage.uoregon.edu/IMAGES/Evaluation_of_Omega_Mag-Bind_TotalPure_NGS_Beads_MWeitzman_April2018.pdf
- Wieser A, Schneider L, Jung J, & Schubert S (2012) MALDI-TOF MS in microbiological diagnostics-identification of microorganisms and beyond (mini review). *Applied Microbiology and Biotechnology* 93:965–974.
- Yang YW, Chen MK, Yang BY, Huang XJ, Zhang XR, He LQ, Zhang J, & Hua Z C (2015) Use of 16S rRNA gene-targeted group-specific primers for real-time PCR analysis of predominant bacteria in mouse feces. *Applied and Environmental Microbiology* 81:6749–6756.
- Zhou Z, Alikhan NF, Mohamed K, Fan Y, Group AS, Achtman M (2020) The Enterobase user's guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia core genomic diversity. *Genome Research* 30:138-152
- Zhou Z, Alikhan N-F, Sergeant MJ, Luhmann N, Vaz C, Francisco AP, Carriço JA, Achtman M (2018) Grapetree: Visualization of core genomic relationships among 100,000 bacterial pathogens. *Genome Research* 28:1395-1404.

APPENDIX

Bioinformatics Script:

- BASH Commands: #Create

SRA environment conda create -n sraenv

python=3.6 #Activate environment conda

activate sraenv #Install SRA-tools conda install

-c bioconda sra-tools #Check installation

worked fastq-dump --help

#Split files using accession number

fastq-dump --split-files --gzip

SRR13651719 fastq-dump --split-files

--gzip SRR13651717

#Install fastp conda install -c bioconda fastp #Run fastp

with your accession number fastp -i

SRR13651719_1.fastq.gz -I SRR13651719_2.fastq.gz -

o out1719.R1.fq.gz -O out1719.R2.2fq.gz fastp -i

SRR13651717_1.fastq.gz -I SRR13651717_2.fastq.gz -

o out1717.R1.fq.gz -O out1717.R2.2fq.gz

- Run the out file in DADA2 pipeline/ ccMetagen website.

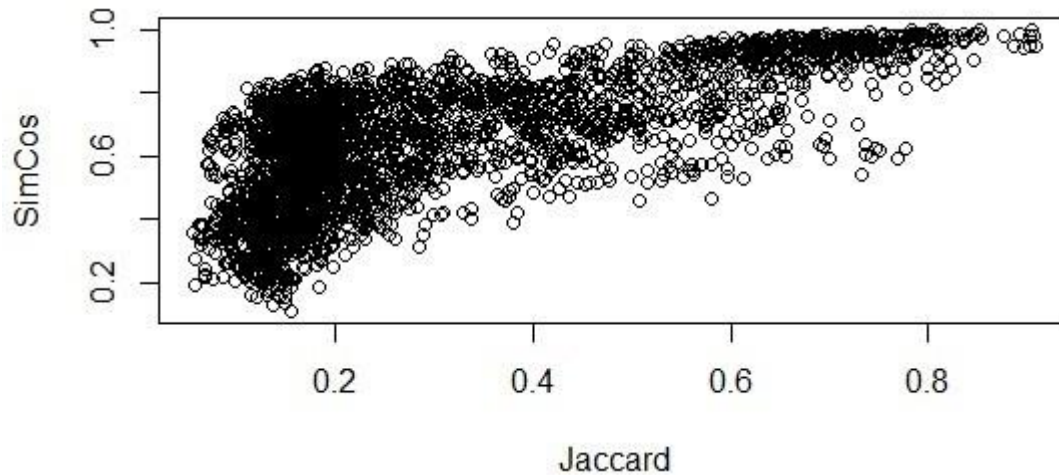
- o DADA2 Pipeline:

file:///C:/Users/akshi/Documents/Spring%2021/BIOL%205334/Results/D
ada2.ht ml

- o ccMetagen:

<https://cge.cbs.dtu.dk/services/CCMetagen-1.2/>

- E. coli MALDI-TOF Jaccard coefficient:



- E.coli MADLI-TOF Script:

<file:///C:/Users/akshi/Documents/E.coli-try2.html>

G-Block sequence:

G-Block sequence as positive control:

>mcr123gBlock

CCAGATAGCTGAACATACACGGCACAGAATACGCCGTCGATGTGCCGCACGATGTGA
CAT

TGCTAAAATTGGTCACGCCATCGATCTTGCAAGCTGTGGGAAAGTATCGCGCTCATA
GC

CATTGAAGCTGACATGATCGGCGCGTGCCGTCTACCGACGACGAACACCACTAGGC
GTG

GCTTACGCATATCAGGCTTGTTGCTTGTACCGCGTCTTTGGCGTGGCAAACCGACCA

AGCCGAGTCTAAGGACTTGATGAATTTGGCGTTTTTTGTGCGAATTATCGGGCTTGGC

GT

GTTGCCAAGTGTGTTGGTCGCAGTTGCCAAAGTCAAACCTCCAGCGTGAGATTGTTCC

AGCCAATTTTCGTTAATAGTACCGTTAAATACGTTTACAATCGTTATCTTGCTGAACCA
 AT
 CCCATTTACAACCTTTAGGTGATGATGCAAAACGGGATACTAATCAAAGTAAGCCCAC
 GTT
 GATGTTTCTGGTCGTTGGTGAAACCGCTCGTGG

iTecan plate reader parameters:

Table: Software parameters for 96 well plate reader

Application: Tecan i-control		Tecan i-control , 1.12.4.0	
Device: infinite 200		Serial number: 911007534	
Firmware: V_2.11_04/08_InfiniTe (Apr 4 2008/14.37.11)		MAI, V_2.11_04/08_InfiniTe (Apr 4 2008/14.37.11)	
Date:	2/1/2022		
Time:	11:08:52 AM		
System		B3506-44487	
User		UHCL\lamontagne	
Plate		Greiner 96 Flat Bottom Black Polystyrene Cat. No.: 655079/655086/655077/655076 [GRE96fb_chimney.pdf]	
Label: PicoGreen dsDNA			
Mode		Fluorescence Top Reading	
Excitation Wavelength		350	nm
Emission Wavelength		460	nm
Excitation Bandwidth		9	nm
Emission Bandwidth		20	nm

Gain		121	Optimal (100%)
Number of Flashes		25	
Integration Time		20	µs
Lag Time		0	µs
Settle Time		0	ms
Part of Plate		A1-D9	
Start Time:	2/1/2022 11:09:06 AM		
	Temperature: 24.8 °C		

MTU table from R-script combined with Bruker ID and score:

	tryE	noiseE	signal	SNR	peaks	MTU	Bruker ID	Bruker Sco	Source
3L	1	1.98E-05	0.000723	36.59308	23	1	Escherichia coli	2.16	seal
DE17	19	1.93E-05	0.000952	49.33034	41	2	Escherichia coli	2.39	dog
L20	27	1.70E-05	0.001201	70.76414	37	2	Escherichia coli	2.39	seal
MB13	54	1.85E-05	0.001129	61.08126	34	2	Escherichia coli	2.39	seal
BB21	2	1.89E-05	0.000926	49.0007	40	2	Escherichia coli	2.42	seal
LE27	45	1.49E-05	0.001261	84.43616	43	2	Escherichia coli	2.43	sewage
DE20	22	1.95E-05	0.00121	62.11994	33	2	Escherichia coli	2.44	dog
DE04	6	2.35E-05	0.000962	40.99852	31	3	Escherichia coli	2.19	dog
SR27	60	1.75E-05	0.001023	58.49031	32	4	Escherichia coli	2.21	OWTS
DE06									
LE24									
LE30									
SR46									
SR40									
SR34									
SR31									
LE32									
ST28									
TA03									

TA01 TA04 ST21 SR44 SR49	8	2.49E-05	0.000832	33.39611	32	4	Escherichia coli	2.35	dog
	42	1.91E-05	0.000939	49.08212	36	4	Escherichia coli	2.38	sewage
	48	2.17E-05	0.000954	43.94273	32	6	Escherichia coli	2.08	sewage
	78	1.46E-05	0.000861	58.97502	43	8	Escherichia coli	2.35	sewage
	73	2.02E-05	0.000903	44.75576	34	10	Escherichia coli	2.25	OWTS
	67	1.28E-05	0.001125	87.61186	38	10	Escherichia coli	2.38	OWTS
	64	1.39E-05	0.001098	78.85287	36	12	Escherichia coli	2.18	OWTS
	50	1.81E-05	0.000895	49.54762	33	12	Escherichia coli	2.35	sewage
							Escherichia coli	2.36	seal
							Escherichia coli	2.38	seal
							Escherichia coli	2.41	seal
							Escherichia coli	2.42	seal
							Escherichia coli	2.42	seal
							Escherichia coli		OWTS
							Escherichia coli		sewage

Kirby-Bauer results for sewage isolates:

Supplementary table: Antibiotic resistance data from different organisms

with the zone of inhibition

Organism	Isolate ID	Gram Strain	Antibiotic	Zone (mm)	Interpretatio n
Enterococcus hirae	Be02	Positive	Chloramphenicol	19	Susceptible
			Ampicillin	0	Resistant
			Tetracycline	24	Susceptible

			Gentamicin	11	Resistant
			Erythromycin	6	Resistant
			Streptomycin	0	Resistant
<i>Aeromonas veronii</i>	Sw75	Negative	Chloramphenicol	18	Susceptible
			Ampicillin	0	Resistant
			Tetracycline	27	Susceptible
			Gentamicin	24	Susceptible
			Erythromycin	0	Resistant
			Streptomycin	22	Susceptible
<i>Paracoccus denitrificans</i>	Sw80	Negative	Chloramphenicol	17	Intermediate
			Ampicillin	0	Resistant
			Tetracycline	13	Intermediate
			Gentamicin	32	Susceptible
			Erythromycin	7	Resistant
			Streptomycin	22	Susceptible
<i>Enterobacter cloacae</i>	Sw59	Positive	Chloramphenicol	12	Resistant
			Ampicillin	0	Resistant
			Tetracycline	16	Intermediate
			Gentamicin	25	Susceptible
			Erythromycin	6	Resistant
			Streptomycin	15	Intermediate
<i>Pseudomonas putida</i>	Sw56	Gram negative	Chloramphenicol	18	Susceptible
			Ampicillin	0	Resistant
			Tetracycline	17	Intermediate
			Gentamicin	20	Susceptible

			Erythromycin	8	Resistant
			Streptomycin	17	Intermediate
Escherichia coli	Sw6	negative	Chloramphenicol	14	Intermediate
			Ampicillin	0	Resistant
			Tetracycline	17	Intermediate
			Gentamicin	22	Susceptible
			Erythromycin	10	Resistant
			Streptomycin	17	Intermediate
Raoultella ornithinolytica	Sw1	negative	Chloramphenicol	19	Intermediate
			Ampicillin	0	Resistant
			Tetracycline	21	Susceptible
			Gentamicin	16	Susceptible
			Erythromycin	7	Resistant
			Streptomycin	16	Intermediate