APPLICATION OF MATRIX-ASSISTED LASER DESORPTION/IONIZATION TIME-OF-FLIGHT MASS SPECTROMETRY (MALDI-TOF MS) FOR DIFFERENTIATING SOURCES OF FECAL POLLUTION

by

Caroline Kmiecik, BS

THESIS

Presented to the Faculty of

The University of Houston-Clear Lake

In Partial Fulfillment

Of the Requirements

For the Degree

MASTER OF SCIENCE

in Biological Sciences

THE UNIVERSITY OF HOUSTON-CLEAR LAKE

MAY, 2022

APPLICATION OF MATRIX-ASSISTED LASER DESORPTION/IONIZATION TIME-OF-FLIGHT MASS SPECTROMETRY (MALDI-TOF MS) FOR DIFFERENTIATING SOURCES OF FECAL POLLUTION

by

Caroline Kmiecik

APPROVED BY

_____
Michael G. LaMontagne, PhD., Chair

_____
Lory Santiago-Vazquez, PhD., Committee Member

_____
Cindy Howard, PhD., Committee Member

APPROVED/RECEIVED BY THE COLLEGE OF SCIENCE AND ENGINEERING:

_____
David Garrison, PhD., Associate Dean

_____
Miguel A. Gonzalez, PhD., Dean

Dedication

It is with genuine gratitude that I dedicate my research to my thesis advisor, Dr. LaMontagne. Dr. LaMontagne first opened the door of opportunity in research for me as an undergraduate in 2018, and allowed me to consistently conduct research throughout the completion of my Master's degree. The level of responsibility he has given me and the trust he has had in me throughout my time working with him has inspired a level of confidence in my research abilities that I would never have thought possible. The years I have spent conducting research in microbiology at UHCL has been the time of my life and I feel honored to have had Dr. LaMontagne as a mentor throughout my academic career. Thank you for everything, Dr. LaMontagne.

Acknowledgements

I would like to thank the University of Houston – Clear Lake for the opportunity to conduct independent research and complete my Master's degree. I would like to thank my committee members, Dr. Santiago and Dr. Howard, who helped foster my interest in biotechnology and environmental research, and provided encouragement and support in my work. I would like to especially thank my advisor, Dr. LaMontagne, for the knowledge, guidance and encouragement he provided throughout my research.

I would like to thank both Moody Gardens and the Dallas Salmon Wastewater Treatment Plant for providing the harbor seal scat and sewage samples, respectively, that allowed me to conduct my thesis research.

I would also like to send my gratitude to all the current and previous microbiology independent study students who have provided me with support throughout my research. In particular, I would like to thank Akshita Mirani and Farnaz Mazhari for the help and encouragement they both gave me throughout my graduate career.

ABSTRACT

APPLICATION OF MATRIX-ASSISTED LASER DESORPTION/IONIZATION TIME-OF-FLIGHT MASS SPECTROMETRY (MALDI-TOF MS) FOR DIFFERENTIATING SOURCES OF FECAL POLLUTION

Caroline Kmiecik
University of Houston-Clear Lake, 2022

Thesis Chair: Michael G. LaMontagne, PhD.

Since the passage of the Marine Mammal Protection Act in 1973, pinniped populations in coastal waters of the United States have increased exponentially. These high populations of large mammals could contribute to fecal contamination of recreational waters. *Enterococci* species counts are used to assess the degree of fecal contamination and elevated counts of this fecal indicator bacteria (FIB) force managers to close beaches; however, contribution of pinnipeds to high FIB counts is not known. This may reflect the high cost of methods of tracking the source of microbial contamination. Matrix Assisted Laser Desorption Ionization – Time of Flight Mass Spectrometry (MALDI-TOF MS) is a time and cost-effective way to identify bacteria through protein mass spectra analysis. MALDI-TOF can distinguish strains of bacteria of the same species but has not been evaluated as a tool for tracking Enterococci isolated from pinnipeds. In this study,

Enterococci were isolated from fresh seal scat samples taken from a pen housing harbor seals. Libraries of Enterococci were also generated from dog scat and a sample from a wastewater treatment plant. These isolates were identified with a MALDI-TOF system and confirmed by 16S rRNA gene sequencing. Cluster analysis of mass spectra generated by MALDI-TOF grouped 22 *E. faecalis* isolates clustered into four coherent MALDI-TOF taxonomic units (MTUs), with two MTUs being harbor seal-specific. Phylogenetic analysis of *E. faecalis* 16S rRNA sequences from harbor seal and dog scat showed that sequences from different mammalian and non-mammalian sources clustered together, confirming that 16S rRNA sequencing is unable to differentiate sources of *E. faecalis* contamination. The distinct source-dependent MALDI-TOF MS clusters suggest that MALDI-TOF MS may be a valuable tool in microbial source tracking.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I:

INTRODUCTION

Pinniped Population of Cape Cod, Massachusetts

Harbor seals (*Phoca vitulina*) and Gray seals (*Halichoerus grypus*) are the two most common pinnipeds found year-round in the Cape Cod area of Massachusetts (Sette et al. 2020). Massachusetts and Maine paid bounties on harbor and gray seals from 1888 until 1962 in an effort to increase fishing yields (Lelli et al. 2009). Lelli et al. (2009) estimated that approximately 72,000-135,000 seals were killed during this period, leading to a decrease in the seal populations of Cape Cod, MA to near extinction.

The seal population rebounded following the Marine Mammal Protection Act of 1972. In the last 30 years, surveys of the Cape Cod area have reported approximations of seal populations at 2,035 in 1994 (NOAA, 1994), 15,700 in 2011 (NMFS, 2011), and between 30,000 and 50,000 in 2017 (Moxley et al. 2017). A 2018 survey of the New England region conducted by the National Oceanic and Atmospheric Administration (NOAA) suggested a collective population of 97,000 to 107,000 seals and an annual population growth rate of 4-5% (NOAA 2019).

Harbor Seals (*P. vitulina*)

Harbor seals belong to the *Phocidae* "true seal" family of the mammalian order Carnivora. Harbor seals are carnivorous marine mammals that primarily eat groundfish (NOAA, 2019). Harbor seals are large marine mammals, weighing 180 to 285 pounds and having a length of five to six feet as adults. Harbor seals haul-out in large groups (Figure 1) on rocks and beaches in intertidal zones for resting, foraging, and thermoregulation (Murray 2008).

Figure 1. Harbor seals (Phoca vitulina) hauled out in an isolated area in Chatham, MA. Photo credit: NOAA Fisheries

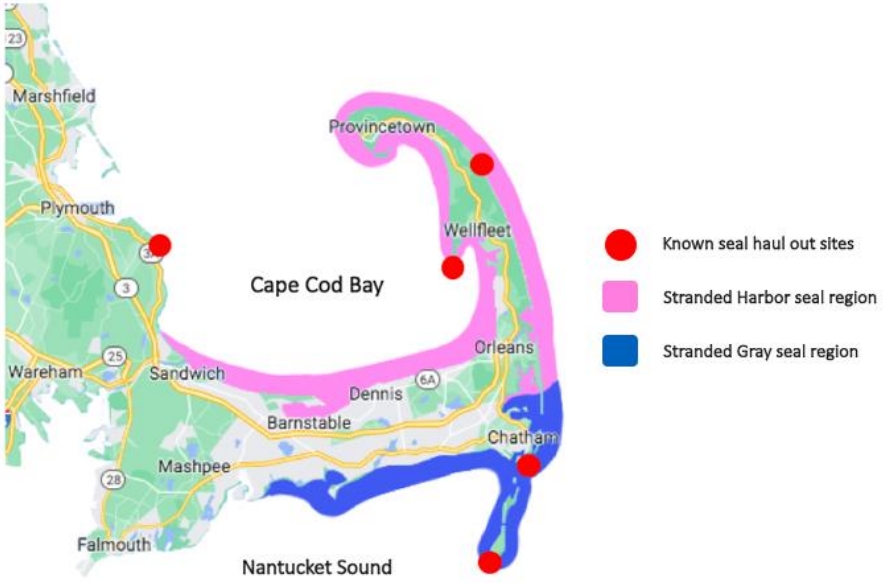

Figure 2. Map of oberserved harbor and gray seal activity recorded between 1999 and 2012, including regions of seal strandings (Frungillo 2014) and known seal haul out sites (Pace et al. 2019).

Ecological Significance of Harbor Seals.

As a top predator, changes in harbor seal abundance, behaviors and health can serve as important indicators of ecosystem change, making them an essential sentinel species (Trites 1997). The turnover and recycling of nutrients is especially crucial role of harbor seals. When harbor seals eat in herds, they release massive ammounts of nutrients (Trites 1997). These nutrients support plankton(Trites 1997).

Though fecal matter from seals provide nutrients to plankton, the increasing numbers of seals and pinnipeds in the wild have contributed to increased fecal contamination of beaches (Costalago et al. 2019). Fecal bacteria are often found in recreational swimming waters near seal habitats (Steele et al. 2018, Staley et al. 2018); however, the level of fecal contribution from wild seals is mostly unknown.

Fecal Contamination in Recreational Waters

A 2021 report by the Environment Massachusetts Research and Policy Center identified 264 beaches out of 556 beaches tested in Massachusetts with unsafe levels of fecal contamination (Weissman and Rumpler, 2021). Of these beaches, 29 were found to have unsafe levels of fecal bacteria on at least 25% of the days tested (Weissman and Rumpler, 2021). These unsafe fecal bacteria levels have resulted in widespread beach closures across Massachusetts. Beach closures resulting from fecal contamination have been a persistent issue for Massachusetts for several years, primarily due to the inability to identify specific sources contributing to the increased contamination.

Microbial Source Tracking

Microbial source tracking is a tool used to differentiate human and non-human sources of fecal pollution, and to identify specific sources of fecal contamination. Microbial source tracking (MST) employs indicator microorganisms to predict the presence of fecal pollution associated with pathogens (Scott et al. 2002).

Indications Fecal Contamination

FIB are found in the intestinal tracts of warm-blooded mammals and are excreted through fecal waste. As high concentrations of FIB are indicative of fecal contamination, which suggests the presence of pathogenic microorganisms and viruses (Scott et al. 2002). Monitoring FIB levels is an essential method in assessing water quality (Gómez-Doñate et al. 2016) as unintentional ingestion of waterborne pathogens may cause a variety of diseases in swimmers. Mild gastroenteritis is the most common of these diseases (Arnold et al. 2016); however, exposure to fecal-contaminated waters can be dangerous for the immunocompromised.

There are four groups of bacteria considered to be fecal indicators: Total coliforms, fecal coliforms, *E. coli*, and *Enterococcus* spp. (Scott et al. 2002). *E. coli* and *Enterococcus* spp. are the most widely used indicators for monitoring fecal pollution as they are abundant in the guts of warm-blooded organisms, have similar survival rates as some pathogens (Scott et al. 2002).

Enterococci

Enterococcus is a large genus of Gram-positive lactic acid bacteria with both clinical and environmental relevance (Devriese et al. 2006). *Enterococci* are facultative anaerobes; this allows Enterococci to exist in a wide variety of environments (Fisher and Phillips, 2009). Enterococci are unable to form endospores; however, they are able to survive in a diverse range of environmental conditions, including extreme temperature (10 – 45 °C), high sodium chloride concentrations (6.5% NaCl), in the presence of bile salts (40%), and wide pH ranges (4.8 –9.6) (Fisher and Phillips, 2009). *Enterococci* are typically found in the intestines of mammals; however, Enterococci are also ubiquitious in nature and can be found free-living in a variety of environments (Manero and Blanch 1999).

Enterococci became the U.S. federal standard for water quality in public marine environments in 2004 (Federal Register 2004). This is partly due to the presence of Enterococci having a higher correlation than fecal coliforms to many human pathogens often found in sewage and water (Jin et al. 2004).

Current Fecal Indicator Limitations

FIB are not limited to humans but also exist in the intestinal tract of many other warm-blooded animals (Orskov 1981, Fiksdal et al. 1985). Though efficient in detecting the fecal contamination, the presence of FIB alone, such as Enterococci, does not provide clear information regarding the specific source of fecal pollution. Enterococci are ubiquitous in nature and can be found in varying concentrations depending on source (Table 1).

Table 1. *Average concentration of Enterococci found in various sources contributing to recreational water contamination*

Source	Enterococci Concentration	Study
Soil	103 CFU/g	Byappanahalli et al. (2012)
Urban Runoff	103 CFU/100 mL-1	Reeves et al. (2004)
Sand	1 - 104 CFU/g	Halliday and Gast (2011)
Bather Shedding	10^6 CFU/Person	Elmir et al. (2007)
Seal Scat	1.21 × 102 - 3.10 × 10^5 CFU/g	Lisle et al. (2004)
Human scat	10^6 CFU/g	Layton et al. (2006)
Dog Scat	104 - 108 CFU/g	Wright et al. (2009)
Raw Sewage	10^5 CFU/100 mL-1	Ahmed et al. (2008)

Culture-Dependent Techniques

Culture-dependent techniques allow researchers to select for and differentiate bacteria from mixed cultures, as well as identify clinically and environmentally-relevant secondary metabolites. These techniques include *in vitro* cultivation of bacteria using

selective and differential growth media. Selective media promotes growth of specific bacteria by inhibiting the growth of other bacteria, while differential media allows the differentiation of bacterial types.

Though less diversity is seen in culture-dependent techniques than in culture-independent techniques (Vaz-Moreira et al. 2011), previous metagenomic studies that focused on the gut microbiome of seals only identified bacteria to a family level of taxonomic classification (Numberger et al. 2016, Pacheco-Sandoval et al. 2019, Glad et al. 2010). To our knowledge, little-to-no studies have been done on identifying harbor seal scat-isolated Enterococci at a species level through culture-dependent techniques.

In this study, we used 2x Rapid Enterococci ChromoSelect Agar (2x REA) (Sigma 51759) and 2x Enterococci Mixed Media broth (2x EMM) to select for and isolate Enterococci from different sources that contribute to fecal contamination of recreational waters. REA and EMM are both selective and differential media, allowing for easier isolation and detection of Enterococci species based on color change in the media. 2x REA contains chromogenic mixture (X-Glu) as an indicator of Enterococci growth; *Enterococci* contain β-D-glucosidase that cleave X-Glu, resulting in colonies with blue-green pigment (Sigma).

Experimental Subjects

A fresh composite scat sample from captive harbor seals was provided by a local aquarium and used to generate a library of Enterococci isolates. To assess the ability of MALDI-TOF to differentiate fecal indicators from different environmental sources, libraries of Enterococci were also generated from domestic dog scat and a composite sample of human waste from a sewage treatment plant, respectively. These sources were chosen based on accessability and their potential contribution to fecal contamination of recreational marine waters.

Domestic Dog (*Canis lupus familiaris*)

Domestic dogs (*Canpis lupus familiaris*) belong to the *Canidae* family of the mammalian order Carnivora. At least 38% of American households own a dog (AMVA 2018), making canines the most popular pet in the US. Dogs are often taken to recreational beaches; however, dog waste not removed from the environment due to irresponsible pet owner behavior has found to contribute to the levels of Enterococci found in beach sand and water (Wright et al. 2009).

Sewage

Sewage waste delivered to sewage treatment facilities undergo a primary and secondary treatment before being discharged back into water outlets. The primary treatment involves filtration processes that remove large particles and sediments from the waste. The secondary treatment results in the removal of approximately 85% of organic matter and is then disinfected with chlorine prior to discharge (EPA 1998). However, different factors such as heavy rainfall overwhelming the combined sewer overflows (Gorman 1999) can contribute to the unintentional discharge of partly or non-treated sewage into the environment, contributing to fecal contamination of recreational waters.

Microbial Identification

Enterococci isolated from harbor seal scat, domestic dog scat, and sewage were identified through MALDI-TOF MS and 16S rRNA sequencing.

MALDI-TOF MS

Matrix-assisted laser desorption-ionization time of flight mass spectrometry (MALDI-TOF MS) is a method of peptide mass fingerprinting (PMF) that provides fast and accurate identifications of unknown microorganisms (LaMontagne et al. 2021). Species and strain discrimination are possible with MALDI-TOF MS (Wolters et al. 2011, Hettick et al. 2006). Additionally, Mazhari (2021) found that MALDI-TOF

MS provides the same species and strain resolution as seen in Whole Genome Sequencing (WGS).

In the full MALDI-TOF MS protocol, formic acid protein extractions from bacterial isolates are spotted onto conductive steel targets and overlaid with a matrix solution. MALDI-TOF MS is a soft ionization method as the matrix solution used to assist the ionization process protects the molecules from being fragmented by the laser pulses. (Wang et al. 2016).

The MALDI-TOF MS system involves desorption and ionization processes in which the matrix is vaporized by pulses of laser light at a wavelength of 337 nm; this ablates analyte and matrix molecules from the steel target. The analyte molecules becoming ionized through proton transfer with the nearby matrix molecules (Wang et al. 2016), resulting in each analyte ion having a single positive charge. In time of flight (TOF) mass spectrometry, an electric field accelerates these ionized molecules so that each ion has the same kinetic energy; this allows separation of ions by mass as the lighter ions travel faster than heavier ions (Wang et al. 2016).

MALDI-TOF MS can identify bacteria by comparing PMFs with reference spectra databases. Bacteria-specific mass spectra found within a mass range of 2 to 20-kDa (Panda et al. 2014) are typically used for this comparision. The popular MALDI Biotyper reference library consists of 3,893 species across 664 microorganism genera (Bruker 2021), allowing for easy identification for diverse groups of bacteria, including FIB

16S rRNA Sequencing

16S rRNA sequencing is commonly used for phylogenetic analysis of bacteria (Johnson et al. 2019). Comparing 16S rRNA sequences with reference sequences available on public databases allows for the differentiation between bacteria across major

phyla at the genus level (Janda and Abbott 2007). Universal 16S primers can be used to amplify the 16S rRNA sequence extracted from unknown isolates, giving a large yield of 16S rRNA for sequencing and microbial identification. Though 16S rRNA sequencing can identify many genera of bacteria to the species and level, the ability to fully resolve taxonomic identities at these levels is often limited. Limitations to 16S identifications include inadequate references in databases not allowing for proper comparison, as well as many 16S rRNA sequences being partial due to short read sequencing platforms targeting the 16S variable regions. Limiting 16S sequencing to only a few hypervariable regions does not give the taxonomic resolution that would be allowed by sequencing the entire approximately 1,500 base pair 16S rRNA gene (Johnson et al. 2019).

Significance

The persistent issue of fecal pollution in recreational marine waters primarily remains due to the inability to identify the source of fecal contamination (Bernhard and Field, 2000). Identifying sources of fecal contamination is essential for assessing associated human health risks and determining the necessary actions for remediation (Scott et al., 2002). Fecal pollution in marine environments can come from a variety of places, including: human sewage, agricultural livestock runoff, domestic pets, and marine mammals such as harbor seals. As there are so many different elements that can contribute to marine fecal pollution, source identification is crucial.

To our knowledge, this is the only study to date to assess the capability of MALDI-TOF MS to differentiate Enterococci isolated from pinnipeds and Enterococci isolated from different potential sources of recreational water fecal contamination.

2018 Preliminary Data

As pinnipeds are protected by the 1976 Marine Mammal Protection Act, studying seals in the wild can be difficult. To assess whether captive seals could be used to study wild seals, DNA sequences of bacteria isolated from captive harbor seal scat in a 2018 preliminary study was used to construct a phylogenetic tree.

In this 2018 study, DNA was extracted from 180 mg of harbor seal scat using the NucleoSpin DNA Stool kit (Macherey-Nagel, #740472). The extracted DNA was used to amplify the 16S rRNA region using Polymerase Chain Reaction (PCR) with universal 16S primers. Following 16S rRNA sequencing and manual curation of sequences generated from harbor seal scat, a phylogenetic tree was constructed. This tree includes sequences isolated from other marine mammals and terrestrial mammals. Sequences from captive harbor seals were found to be closely related to sequences isolated from wild seals (Figure 3). Sequences from both captive and wild seals clustered closely with other marine mammalian sequences; however, marine mammal sequences did not cluster with sequences from terrestrial mammals. This suggests that captive seals are a convenient, reliable way to study the gut microbiome of wild seals.

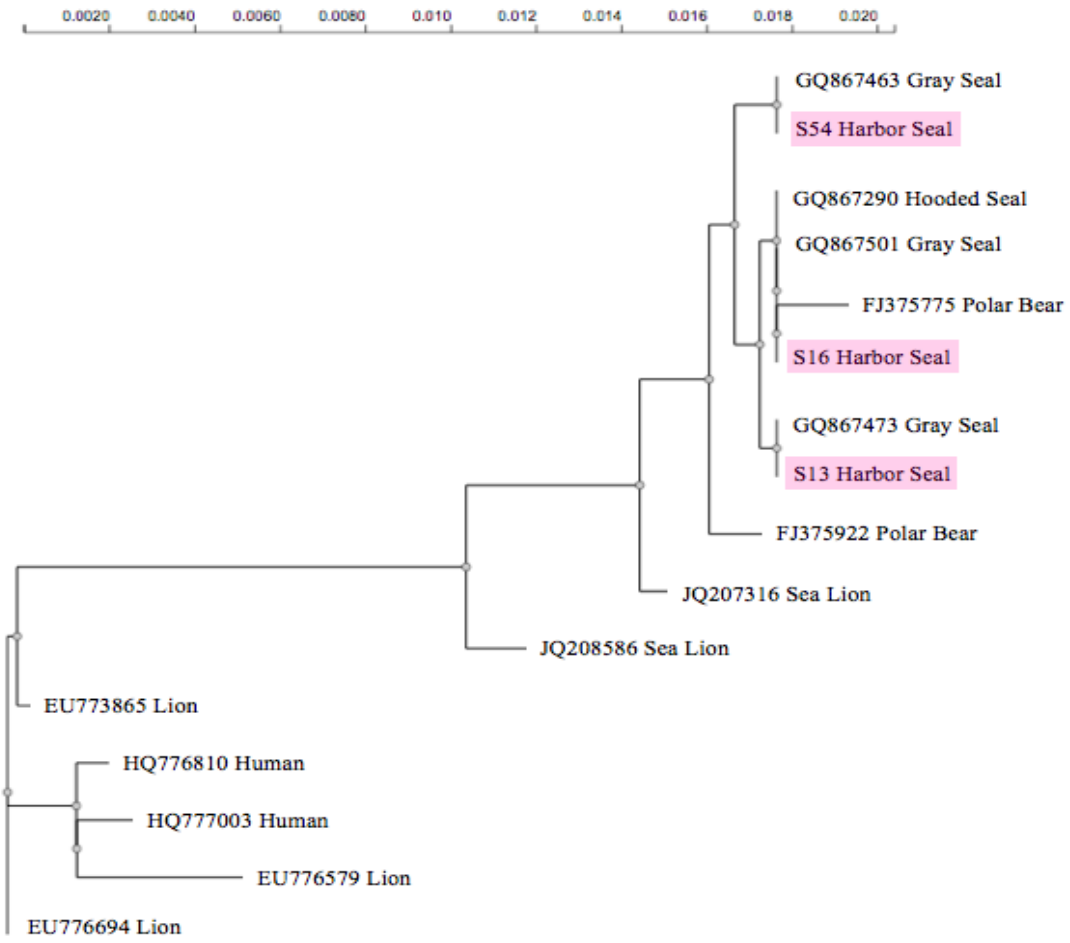

Figure 3. 16S rRNA phylogenetic analysis from 2018 preliminary study showing that gut bacteria isolated from captive seals are closely related to those from wild seals. Highlighted in pink are sequences derived from captive harbor seals.

CHAPTER II:

METHODOLOGY AND MATERIALS

Sample Collection

Six grams of a fresh composite scat sample was provided by a local aquarium housing harbor seals. These captive seals are kept in a 70,000 gallon tank filled with saltwater treated with ozone; temperature is maintained at 15.56°C and salinity at 15 ppt. Fresh domestic dog scat was collected and provided by the owner. A composite sample of human sewage in a 50 mL conical tube was provided by the Dallas Salmon Wastewater Treatment Plant (Houston, TX, USA). All samples collected were transported to the lab on ice and immediately used for serial dilution preparation.

Scat Sample Preparation

One gram of fresh scat was transferred to a sterile 50 mL conical tube and suspended in 20 mL of 1x PBS to make a scat slurry. Slurries were vortexed until homogenous and centrifuged at 10,000 x g for 10 minutes. 15 mL of the resulting supernatant was decanted and discarded. The remaining 5 mL of supernatant was transferred to a new, sterile 50 mL conical tube to be used for serial dilutions. Slurries were serially diluted in 1x PBS from 10^{-1} to 10^{-6} for plating on Enterococci-selective media. The remaining scat slurries were preserved in 10% glycerol and stored at -80℃ to retain viability.

Cultivation of Enterococci

2x Rapid Enterococci ChromoSelect Agar

100 µl of 10^{-4}, 10^{-5}, and 10^{-6} dilutions were transferred to a single quadrant on 2x Rapid Enterococci ChromoSelect Agar (2x REA) and streaked with a sterile swab. The remaining plate was quadrant streaked with sterile loops to isolate colonies of Enterococci. 2x REA plates was incubated at 37℃ for 48 hours. Each dilution was plated

on to three separate plates per source, for a total of 27 plates. Blue colonies indicative of Enterococci growth were restreaked onto 2x REA agar to verify isolation of Enterococci.

2x Enterococci Mixed Media Broth

EMM broth was prepared with 42 mL of 2x EMM base (2 g peptone, 2 g yeast extract, 1 g lactose, 0.54 g potassium dihydrogen phosphate, 2 g sodium citrate in 84 mL of distilled water), 5 mL of 1% sodium azide, 2 mL of 0.88 M sodium carbonate monohydrate, and 1 mL of Trinitrophenyl for a final volume of 50 mL.

A 96-well microplate containing 200 µl of 2x Enterococci Mixed Media (EMM) broth were inoculated with 50 µl of 10^{-4}, 10^{-5}, and 10^{-6} dilutions and incubated at 37℃ for 48 hours. Wells positive for Enterococci growth were restreaked onto 2x REA agar to verify isolation of Enterococci. 100 µl from the well plate was transferred to a quadrant on 2x REA and streaked with a sterile swab. The remaining plate was quadrant streaked with sterile loops to isolate colonies of Enterococci. All isolated colonies were restreaked onto TSA and prepared for MALDI-TOF MS within 24 hours of growth (Table 2).

Table 2. *Isolate codes and the source sample and dilution cultured from.*

Dilution Plated	Harbor Seal	Domestic Dog	Sewage
EMM Well 10^{-6}	ES01-ES04	ED01-ED04	EL01-EL04
EMM Well 10^{-5}	ES05-ES08	ED05-ED08	EL05-EL08
EMM Well 10^{-4}	ES09-ES12	ED09-ED12	EL09-EL12
10^{-6}	ES13-ES24	ED13-ED24	EL13-EL24
10^{-5}	ES25-ES40	ED25-ED40	EL25-EL40
10^{-4}	ES41-ES56	ED41-ED56	EL41-EL56

MALDI-TOF MS

Proteins were extracted from isolates following the Formic Acid/Ethanol Tube Extraction (TE) Method (Bruker). A 10 µl-loopful of subculture from an isolated pure bacterial colony was suspended in 300 µl LC-MS Grade Water (Sigma WX001) and vortexed until homogenous. Bacterial suspensions were treated with 70% ethanol by adding 900 µl of 100% HPLC-MS Ethanol (Sigma 459828) and vortexing thoroughly. Suspensions were centrifuged at 14,500 rpm for two minutes and the supernatant was completely decanted and discarded. Bacterial pellets were allowed to air-dry at room temperature for 10 minutes to ensure all ethanol was removed.

Pellets were resuspended in 50 µl of 70% formic acid (Fisher Scientific #A117-50) and vortexed thoroughly; tubes were allowed to stand for five minutes. 50 µl of 100% acetonitrile (Thermo Scientific #51101) was added to each tube and vortexed thoroughly for 10 seconds. Formic acid extractions were centrifuged at 14,500 rpm for two minutes, and 70 µl of the resulting supernatant was transferred to a sterile 1.5 mL microcentrifuge tube. Protein extractions were stored at -20℃ prior to steel target spotting.

Prior to spotting protein extractions, a 384-spot steel target plate was cleaned with 70% EtOH and 80% Trifluoroacetic acid (TFA). The target was rinsed with distilled water and allowed to dry completely.

One microliter of protein extraction was placed onto a spot on the 384-spot steel target plate and allowed to air dry. The supernatant spots were then overlaid with 1 µl of matrix solution (5 mg Bruker HCCA/mL) and allowed to air-dry. The steel targets were then packaged on ice and shipped overnight to the Proteomics and Mass Spectrometry Core Facility at the Huck Institute (The Pennsylvania State University, University Park, PA, USA) for mass spectrometry analysis.

MALDI-TOF MS Analysis

Identifications of isolates from the Bruker MALDI Biotyper system and their associated reliable-ID scores were received in the form of a PDF report. The reliabilty of isolate identities were determined using score range presented in Table 3 (Bruker 2021).

Mass spectra were analyzed by cluster analysis using custom scripts (LaMontagne et al. 2021) and functions from the R package MALDIquant (Gibb and Strimmer 2012). The script used included two optimization loops that repeatedly sampled random values within a specified range for the following parameters: smoothing, baseline removal, alignment, alignment tolerance, and signal to noise ratio (LaMontagne et al. 2021). The first loop optimized the number of peaks shared between pairs of average mass spectra and calculated Jaccard coefficients. The second optimization loop minimized the overlapping of cosine similarity values between closely and distantly related isolates, respectively (LaMontagne et al. 2021, Strejcek et al. 2018). The script defined MALDI-TOF taxonomic units (MTUs), which correspond to operational taxonomic units, through cluster analysis (LaMontagne et al. 2021). Approximately unbiased probability values and bootstrap probabilities were assigned to clusters with Pvclust (Suzuki and Shimodaira 2006). The script used in this study is presented in an R markdown file (Appendix).

Table 3. *Reliability of Bruker IDs by score range.*

Score Range	Description
2.300 ... 3.000	Highly probable species identification
2.000 ... 2.299	Secure genus identification, probable species identification
1.700 ... 1.999	Probable genus identification
0.000 ... 1.699	No reliable identification

DNA Isolation

Representative isolates of *E. faecalis* MTUs that were identified as species-secure using the Bruker MALDI Biotyper database (Biotyper score > 2.3) were identified by 16S rRNA gene sequencing to confirm the identity of MALDI-TOF MS isolates.

DNA was extracted from Enterococci isolates using the MasterPure™ Gram Positive DNA Purification Kit (Lucigen MGP04100). Overnight cultures of Enterococci isolates in 1 mL TSB were pelleted by centrifugation at 16,000 x g for two minutes; the resulting supernatant was decanted and discarded. The pellets were resuspended in 140 µl TE Buffer and vortexed until homogenous. 10 µl TE with 1 µl Ready-Lyse Lysozyme was added to each tube and vortexed to mix. Tubes were incubated overnight in a 37℃ water bath to help facilitate lysis. After 24 hours, 150 µl Gram-Positive Lysis solution with 1 µl Proteinase K (50 µg/µl) was added to lysates in each tube, and vortexed thoroughly to mix. Tubes were incubated at 70℃ for 15 minutes at 300 rpm, and briefly vortexed every five minutes. 2 µl of RNase A (10 µg/µl) was added to each tube and incubated in a 37 ℃ water bath for 30 minutes. 175 µl of MPC Protein Precipitation was added to each tube and vortexed thoroughly for 10 seconds. Cells were pelleted by centrifuging at 4℃ for 10 minutes at 10,000 x g. The resulting supernatant was transferred to new, sterile 1.5 mL microcentrifuge tube. 250 µl of ice-cold isopropanol was added to the supernatant and mixed by inverting tubes 40 times. DNA was pelleted by centrifuging at 4℃ for 10 minutes at 10,000 x g. The isopropanol was decanted and discarded, and DNA pellets were washed twice with 10 µl of 70% ethanol. Extracted DNA was eluted in 35 µl TE buffer and DNA concentration was measured using NanoDrop. DNA extractions were stored at -20℃ prior to amplification with 16S PCR.

16S rRNA PCR

DNA extractions and universal primers were used to amplify the 16S rRNA gene of Enterococci isolates through PCR. Enterococci DNA templates were thawed on ice and flicked into suspension if needed. PCR master mix was prepared in a sterile 1.5 mL microcentrifuge tube using 120 µl of 2x Platinum II Hot-Start PCR Master Mix (Invitrogen #14000012), 100 µl of nuclease-free water (Invitrogen #4387936), 10 µl of 0.1 µM 8F (5'-AGAGTTTGATCCTGGCTCAG-3'), and 10 µl of 0.1 µM 1492r (5'-GGTTACCTTGTTACGACTT-3'), for a final volume of 240 µl.

24 µl of the prepared master mix was aliquoted into ten separate 0.2 mL PCR tubes. One microliter of DNA template was transferred to nine tubes, giving a final reaction volume of 25 µl. The tenth tube was inoculated with 1 µl of nuclease-free water to be used as a negative, no-template control. Thermocycling was performed as described in Table 3.

Table 4. *PCR Thermocycling Conditions*

Step	Temperature	Time
Initial Denaturation	94℃	2 minutes
PCR: 35 Cycles		
Denaturation	98℃	5 seconds
Annealing	60℃	15 seconds
Elongation	60℃	15 seconds
Hold	4℃	∞

PCR products were examined for amplification through agarose gel electrophoresis. A 1% agarose gel was prepared using 1x TBE buffer and GelRed. Six microliters of 1kb DNA ladder (NEB N3200S) was placed in the first well to analyze DNA base pair size. Five microliters of PCR product supplemented with 1 µl of 6x loading dye (NEB B7021), and the no-template control were transferred into the

remaining wells. Gel electrophoresis was run at 75V until the DNA traveled down three quarters of the gel. The gel was then observed using the Gel Doc XR System from Biorad to check for successful amplification. PCR products with amplified DNA were shipped to Lone Star Laboratory (Houston, TX, USA) for 16S rRNA sequencing.

16S rRNA Phylogenetic Analysis

The generated forward and reverse 16S rRNA sequences were manually curated in the software program BioEdit (7.2.5). The reverse sequence was reverse-complemented and pair-wise aligned with the forward sequence. Mismatched nucleotide base pairs were manually curated using the generated 16S spectra data. A consensus sequence was generated from the aligned forward and reverse sequences to be used for phylogenetic analysis. Consensus sequences were used to find reference GenBank sequences using BLAST. Consensus and reference sequences were automatically aligned by CLUSTAL-W. A phylogenetic tree with the multiple sequence alignment was constructed in BioEdit using the DNAmlk DNA Maximum Likelihood program with molecular clock method. The generated tree file was visualized using the web service phylogeny.fr (Dereeper and Guignon et al. 2008).

CHAPTER III:

RESULTS

Cultivation of Enterococci

56 isolated colonies with morphologies consistent with Enterococci were selected per source, for a total of 168 isolates. Upon restreaking onto 2x REA, 144 isolates were found to have morphologies consistent with Enterococci: 53 isolates from seals, 56 isolates from dogs, and 47 isolates from sewage.

2x EMM Broth

Nine sewage isolates restreaked from Enterococci-positive 2x EMM 10^{-4} wells (Figure 4) did not have the blue pigment indicative of Enterococci growth when plated onto 2x REA; however, four additional sewage isolates from the 2x EMM 10^{-4} wells were restreaked for MALDI-TOF analysis, each with morphologies consistent with Enterococci. Enterococci-positive 2x EMM wells from harbor seal scat and dog seal scat plated onto 2x REA, respectively, all had colony morphologies consistent with that of Enterococci (Figure 5).

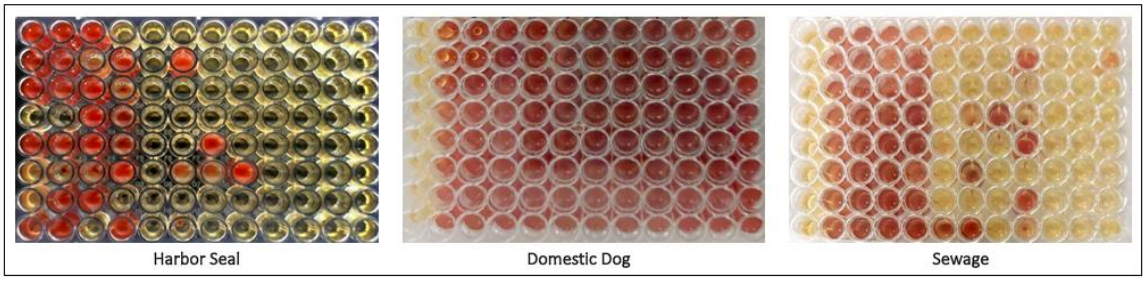

Figure 4. 2x EMM plates with labels indicating inoculation source sample. Red wells are indicative of Enterococci growth. The twelfth column for harbor seal and the first column for domestic dog and sewage are uninoculated wells used as a negative control.

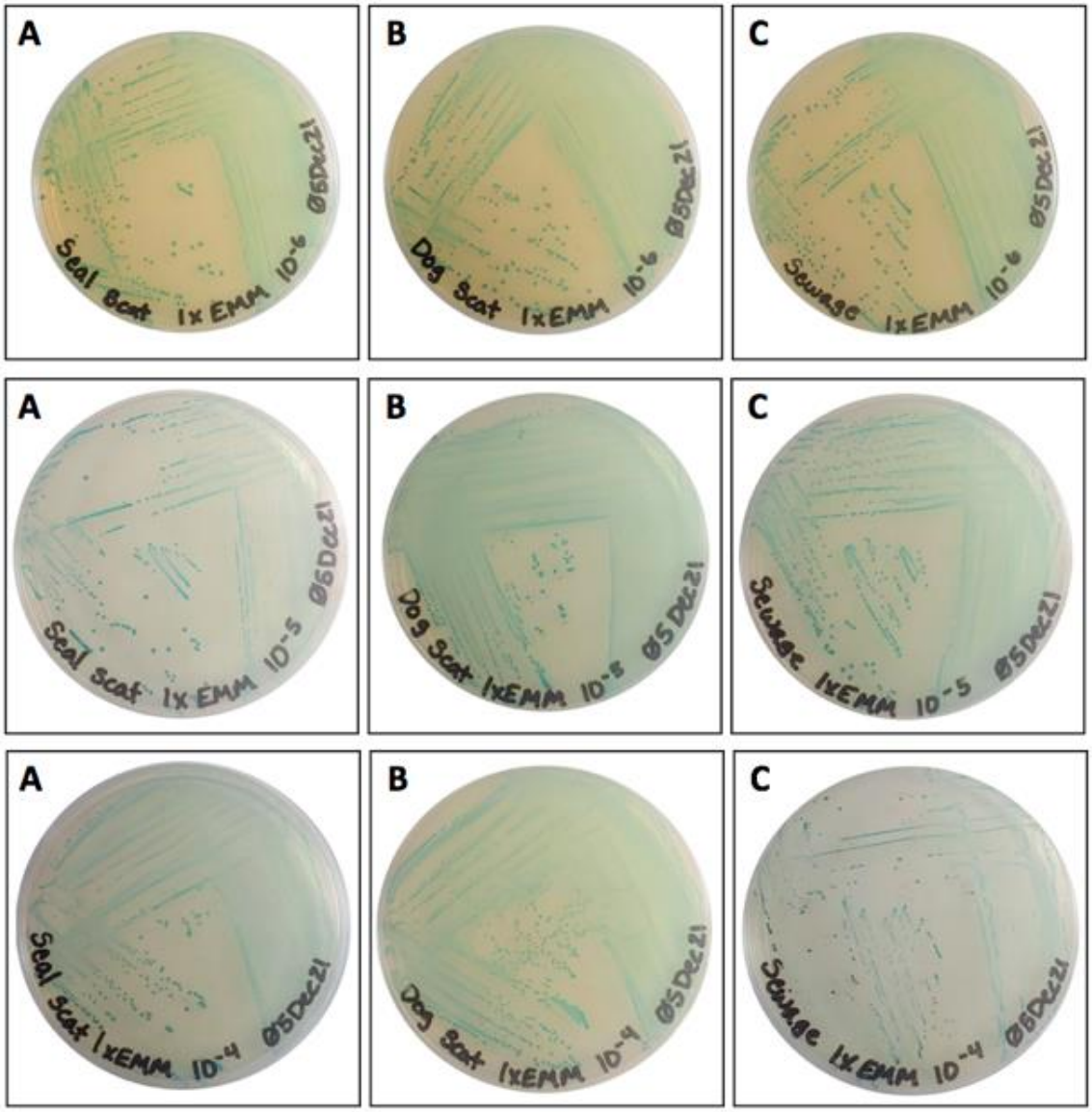

Figure 5. Dilutions from 2x EMM wells plated onto primary 2x REA plates for A) Harbor seal, B) Domestic dog, and C) Sewage.

2x Rapid Enterococci ChromoSelect Agar

Isolates from the 10^{-5} and 10^{-6} dilutions plated onto primary 2x REA plates all had morphologies consistent of Enterococci when restreaked onto secondary 2x REA (Figure 6). Two seal scat isolates and one sewage isolate from their respective 10^{-4} dilution primary 2x REA plates did not have morphologies consistent with that of Enterococci and were not included in MALDI-TOF MS preparations. 42 isolates per source were selected for MALDI-TOF MS analysis, for a total of 126 isolates identified by the Biotyper database.

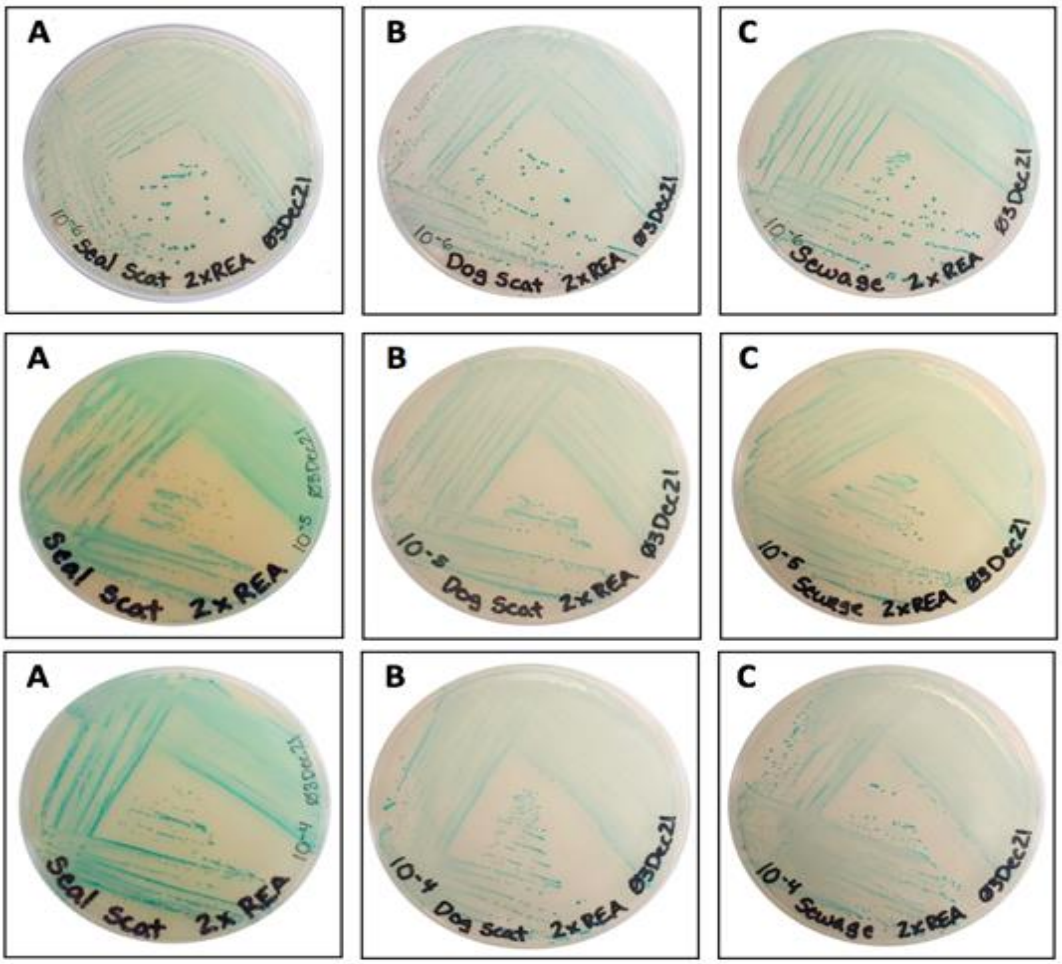

Figure 6. Slurry dilution 2x REA primary plates for A) Harbor seal scat, B) Dog scat, and C) Sewage. 10^{-6}, 10^{-5}, and 10^{-4} dilutions plated onto 2x REA show in first, second, and third row, respectively.

MALDI-TOF MS Identifications

Of the 126 isolates identified via MALDI-TOF MS, 114 were identified as Enterococci, with 54 species-secure isolates, 58 genus-secure isolates, and two genus-probable isolates. Six Enterococci species were identified from the collective source samples: *E. casseliflavus, E. faecium, E. faecalis, E. gallinarum, E. hirae,* and *E. mundtii* (Table 5). Despite having morphologies consistent with that of Enterococci on 2x REA, four isolates cultured from harbor seal 2x EMM 10^{-5} wells (ES04-ES08) were identified as *Carnobacterium maltaromaticum*, and thus excluded from cluster analysis.

Table 5. *Enterococci species identified from each sample source.*

Enterococcus Species	Harbor Seal	Domestic Dog	Sewage
E. casseliflavus	0	8	0
E. faecium	0	0	34
E. faecalis	34	18	1
E. gallinarum	0	1	0
E. hirae	4	0	3
E. mundtii	0	11	0

Enterococci isolated from harbor seal scat were identified as either *E. faecalis* or *E. hirae* (Table 6). Domestic dog scat had the greatest species diversity of Enterococci cultured, with a total of four species isolated. Sewage samples had a total of three Enterococci species isolated, while only two Enterococci species were isolated from harbor seal scat. These results are in partial unison with those of Medeiros et al. (2017), who identified high proportions of *E. faecalis* and the presence of *E. hirae* in South American and Subantarctic fur seals samples through qPCR. However, Medeiros et al. (2017) also detected a high proportion of *E. mundtii* and the presence of *E. gallinarum*

and *E. casseliflavus*, both of which isolated from dog scat (Table 7) but was not isolated from seal scat in this study.

Enterococci isolated from domestic dog scat were identified as *E. faecalis*, *E. casseliflavus*, *E. mundtii, and E. gallinarum* (Table 7). Enterococci isolated from treated sewage samples were identified as *E. faecium*, *E. faecalis*, and *E. hirae* (Table 8). Isolates not identified as Enterococci were removed for cluster analysis.

Table 6. *MALDI-TOF Bruker identifications of Enterococci isolated from Harbor Seals. Scores of ≥2.3 are considered to have a species-secured ID. Scores of 2.0 – 2.29 are considered to have a genus-secured ID*

Isolate	Bruker ID	Score
ES01	*Enterococcus faecalis*	2.09
ES02	*Enterococcus faecalis*	2.19
ES03	*Enterococcus faecalis*	2.22
ES10	*Enterococcus faecalis*	2.05
ES11	*Enterococcus faecalis*	2.23
ES12	*Enterococcus faecalis*	2.18
ES13	*Enterococcus faecalis*	2.30
ES14	*Enterococcus faecalis*	2.08
ES15	*Enterococcus faecalis*	2.15
ES16	*Enterococcus faecalis*	1.97
ES17	*Enterococcus faecalis*	2.23
ES18	*Enterococcus faecalis*	2.29
ES19	*Enterococcus faecalis*	2.22
ES20	*Enterococcus faecalis*	2.32
ES21	*Enterococcus faecalis*	2.22
ES22	*Enterococcus faecalis*	2.17
ES23	*Enterococcus faecalis*	2.14
ES24	*Enterococcus faecalis*	2.30
ES25	*Enterococcus faecalis*	2.31
ES26	*Enterococcus faecalis*	2.25
ES27	*Enterococcus faecalis*	2.32
ES28	*Enterococcus faecalis*	2.32
ES29	*Enterococcus faecalis*	2.31
ES30	*Enterococcus faecalis*	2.38
ES31	*Enterococcus faecalis*	2.38
ES32	*Enterococcus faecalis*	2.34
ES33	*Enterococcus faecalis*	2.26
ES34	*Enterococcus faecalis*	2.21
ES35	*Enterococcus faecalis*	2.36
ES36	*Enterococcus faecalis*	2.18
ES37	*Enterococcus faecalis*	2.26
ES38	*Enterococcus faecalis*	2.19
ES39	*Enterococcus faecalis*	2.22
ES40	*Enterococcus faecalis*	2.25
ES43	*Enterococcus hirae*	2.43
ES44	*Enterococcus hirae*	2.31
ES45	*Enterococcus hirae*	2.43
ES46	*Enterococcus hirae*	2.26

Table 7. *MALDI-TOF Bruker identifications of Enterococci isolated from Domestic dog scat.*

Isolate	Bruker ID	Score
ED01	*Enterococcus casseliflavus*	2.16
ED02	*Enterococcus mundtii*	2.37
ED03	*Enterococcus faecalis*	2.30
ED04	*Enterococcus faecalis*	2.11
ED05	*Enterococcus casseliflavus*	2.18
ED06	*Enterococcus mundtii*	2.38
ED07	*Enterococcus faecalis*	2.34
ED08	*Enterococcus faecalis*	2.31
ED09	*Enterococcus faecalis*	2.34
ED10	*Enterococcus casseliflavus*	2.01
ED11	*Enterococcus faecalis*	2.27
ED12	*Enterococcus mundtii*	2.40
ED13	*Enterococcus faecalis*	2.07
ED14	*Enterococcus mundtii*	2.25
ED15	*Enterococcus casseliflavus*	2.00
ED16	*Enterococcus faecalis*	2.25
ED17	*Enterococcus casseliflavus*	2.05
ED18	*Enterococcus mundtii*	2.39
ED19	*Enterococcus faecalis*	2.35
ED20	*Enterococcus mundtii*	2.30
ED21	*Enterococcus faecalis*	2.13
ED22	*Enterococcus mundtii*	2.30
ED23	*Enterococcus casseliflavus*	2.14
ED24	*Enterococcus mundtii*	2.12
ED25	*Enterococcus casseliflavus*	2.12
ED26	*Enterococcus faecalis*	2.42
ED27	*Enterococcus gallinarum*	2.10
ED28	*Enterococcus casseliflavus*	2.30
ED29	*Enterococcus mundtii*	2.37
ED30	*Enterococcus mundtii*	2.28
ED31	*Enterococcus mundtii*	2.03
ED32	*Enterococcus faecalis*	2.34
ED33	*Enterococcus faecalis*	2.34
ED34	*Enterococcus faecalis*	2.31
ED35	*Enterococcus faecalis*	2.31
ED36	*Enterococcus faecalis*	2.32
ED37	*Enterococcus faecalis*	2.27
ED38	*Enterococcus faecalis*	2.27

Table 8. *MALDI-TOF Bruker identifications of Enterococci isolated from treated-sewage samples.*

Isolate	Bruker ID	Score
EL01	*Enterococcus faecalis*	2.31
EL02	*Enterococcus faecium*	2.52
EL03	*Enterococcus faecium*	2.30
EL04	*Enterococcus faecium*	1.73
EL05	*Enterococcus faecium*	2.37
EL06	*Enterococcus faecium*	2.46
EL07	*Enterococcus faecium*	2.09
EL08	*Enterococcus faecalis*	2.35
EL09	*Enterococcus faecium*	2.20
EL10	*Enterococcus faecium*	2.24
EL11	*Enterococcus faecium*	2.22
EL12	*Enterococcus faecium*	2.24
EL13	*Enterococcus faecium*	2.30
EL14	*Enterococcus faecium*	2.35
EL16	*Enterococcus hirae*	2.38
EL17	*Enterococcus faecium*	2.26
EL18	*Enterococcus faecium*	2.19
EL19	*Enterococcus hirae*	2.44
EL20	*Enterococcus faecium*	2.05
EL21	*Enterococcus faecium*	2.30
EL22	*Enterococcus faecalis*	2.25
EL23	*Enterococcus faecium*	2.35
EL25	*Enterococcus faecium*	2.44
EL26	*Enterococcus faecium*	2.25
EL27	*Enterococcus faecium*	2.41
EL28	*Enterococcus faecium*	2.19
EL29	*Enterococcus faecium*	2.18
EL30	*Enterococcus faecium*	2.45
EL31	*Enterococcus faecium*	2.22
EL32	*Enterococcus faecium*	2.26
EL33	*Enterococcus faecium*	2.41
EL34	*Enterococcus hirae*	2.42
EL35	*Enterococcus faecium*	2.24
EL36	*Enterococcus faecium*	2.43
EL37	*Enterococcus faecium*	2.22
EL38	*Enterococcus faecium*	2.18
EL39	*Enterococcus faecium*	2.24
EL40	*Enterococcus faecium*	2.35

MALDI-TOF Cluster Analyses

All Enterococci Isolated

Enterococci isolates from each source clustered into 17 separate MTUs (Figure 7). Each MTU was specific to a single source, except for MTU03 and MTU05, which had isolates from seal and dog scat. Three MTUs were specific to harbor seals, two were specific to domestic dogs, and nine were specific to sewage (Table 9).

The source-specific MTU clusters are likely due to the difference in *Enterococcus* species diversity isolated from each source; however, there were three harbor seal-specific MTUs that did not cluster with isolates of the same species from dog scat and sewage samples. MTU15 had 14 harbor seal-isolated *E. faecalis*, while MTU16 and MTU17 had harbor seal-isolated *E. hirae*.

Approximately two-thirds (3095/4465) of the pairwise comparisons showed Jaccard similarity coefficients of less than 0.2 (Figure 8). Cosine similarity for these pairwise comparisons ranged from 0.068 to 0.497. Approximately one-third (1370/4465) of the pairwise comparisons showed Jaccard similarity coefficients greater than or equal to 0.2. Cosine similarity for these pairwise comparisons ranged from 0.500 to 0.833. Jaccard similarities of 0.2 corresponded to cosine similarities of 0.50. Isolates clustered within MTUs appeared coherent, as they were consistent with the species identified by the Bruker MALDI Biotyper reference library.

Consistent mass-to-charge ratios (m/z) were observed across the *Enterococcus* genus with distinct peaks at 4411.953, 4758.186, 5095.4, 7272.734, and 9543.53. Peaks at 7272.734 and 9543.53 were the most intense amongst these consistent m/z values. Several peaks were found to be shared amongst *Enterococcus* species as well, with distinct peaks at 8971.889 and 9054.732 observed in all *E. faecium*, *E. mundtii*, and *E.*

faecalis mass spectra; however, these peaks were not found in *E. hirae* or *E. casseliflavus.*

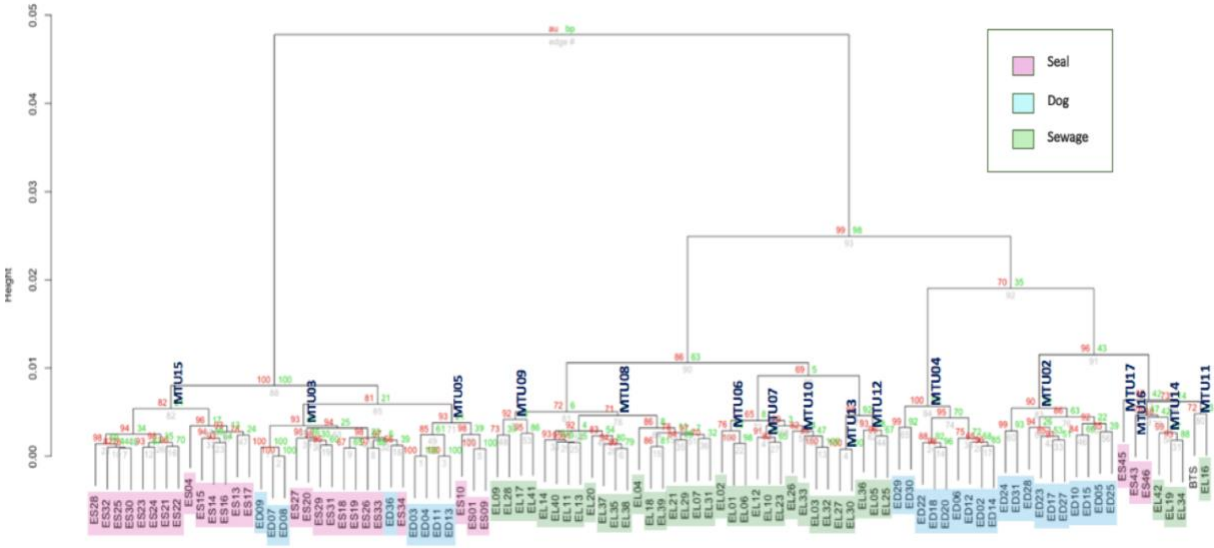

Figure 7. Hierarchical clustering of mass spectra for Enterococci isolated from each sample source. Pink, blue, and green boxes represent seal, dog, and sewage isolates, respectively. MTUs are indicated at each node.

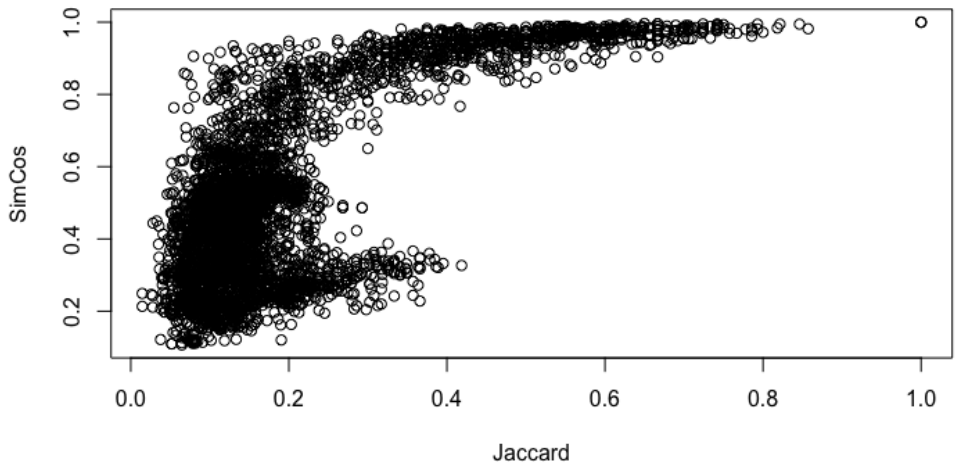

Figure 8. Cosine versus Jaccard similarities for pairwise comparisons of mass spectra from all Enterococci isolates.

Table 9. *MTUs defined through cluster analysis showing the number of isolates from each source per MTU. MTU boxes highlighted indicate MTUs with isolates from a single source. Blue, green, and pink boxes indicate MTUs with dog, sewage, and seal isolates, respectively.*

MTU	Harbor Seal	Domestic Dog	Sewage
MTU02	0	10	0
MTU03	9	4	0
MTU04	0	9	0
MTU05	3	4	0
MTU06	0	0	3
MTU07	0	0	3
MTU08	0	0	15
MTU09	0	0	4
MTU10	0	0	4
MTU11	0	0	1
MTU12	0	0	3
MTU13	0	0	2
MTU14	0	0	3
MTU15	14	0	0
MTU16	2	0	0
MTU17	1	0	0

Species-Secure *E. faecalis*

22 species-secure *E. faecalis* isolates (11 from seal, 11 from dog) were used for cluster analysis. *E. faecalis* from seal and dog scat clustered into four MTUs (Figure 9). Two of these MTUs (MTU01 and MTU02) had isolates from both seals and dogs; however, MTU03 and MTU04 only had seal-isolated Enterococci (Table 10). Seal isolate ES35 and dog isolate ED08 in MTU02 clustered as sister taxa with an AU value of 91. Though MTU02 had both seal and dog isolates, the two clusters seen within this MTU suggest that isolates from the same source still cluster more closely together.

Enterococci isolates in the seal-specific MTUs (MTU03 and MTU04), as well as the seal-dog mixed MTU01 had a best match pattern to the strain *Enterococcus faecalis* DSM 6134 DSM. Enterococci isolates in the seal-dog mixed MTU02 had a best match pattern to the strain *Enterococcus faecalis* DSM 20409 DSM.

Approximately 14.76% (31/210) of the pairwise comparisons showed Jaccard similarity coefficients of less than 0.2 (Figure 10). Cosine similarity for these pairwise comparisons ranged from 0.413 to 0.496. Approximately 85.24% (179/210) of the pairwise comparisons showed Jaccard similarity coefficients greater than or equal to 0.2. Cosine similarity for these pairwise comparisons ranged from 0.502 to 0.751. Jaccard similarities of 0.2 corresponded to cosine similarities of 0.50.

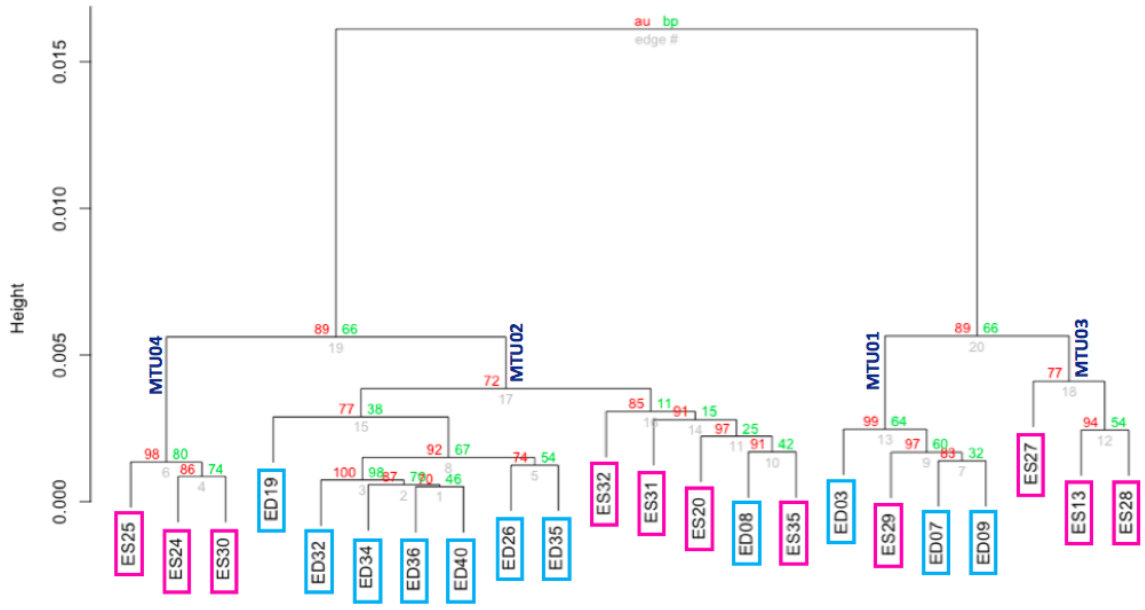

Figure 9. Hierarchical clustering of mass spectra for isolates from seal and dog reliably identified as E. faecalis (Bruker score >2.3) by the Biotyper system. Pink and blue boxes represent seal and dog isolates, respectively. MTUs are indicated at each node.

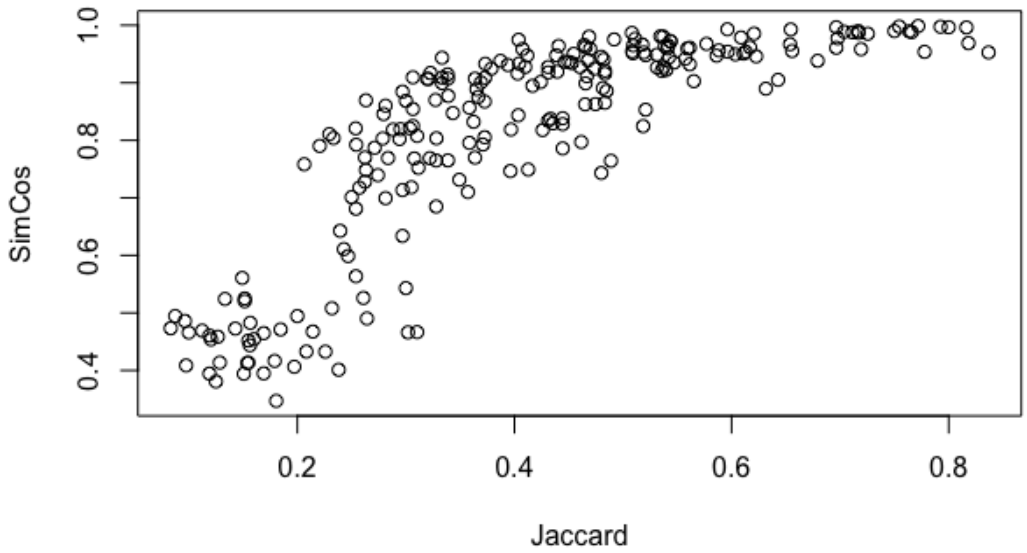

Figure 10. Cosine versus Jaccard similarities for pairwise comparisons of mass spectra from seal and dog species-secure E. faecalis isolates.

Table 10. *MTUs defined through cluster analysis of species-secure Enterococci isolates showing the isolates from seal and dog sources per MTU. Red boxes indicate MTUs that had seal-isolated Enterococci only.*

MTU	Harbor Seal	Domestic Dog
MTU01	(1): ES29	(3): ED03, ED07, ED09
MTU02	(4): ES20, ES31, ES32, ES35	(8): ED08, ED19, ED26, ED32, ED34, ED35, ED36, ED40
MTU03	(3): ES13, ES27, ES28	None
MTU04	(3): ES24, ES25, ES30	None

16S rRNA Sequencing

PCR Amplification of *Enterococci* 16S rRNA

DNA amplification of harbor seal *E. faecalis* isolates ES25, ES27, ES29 and ES32, and dog *E. faecalis* isolates ED08, ED09 and ED36 was assessed through gel electrophoresis. Each PCR product had a DNA band size of approximately 1,500 base pair (Figure 11), which is consistent with the expected PCR product size and known length of the 16S rRNA gene (Johnson et al. 2019).

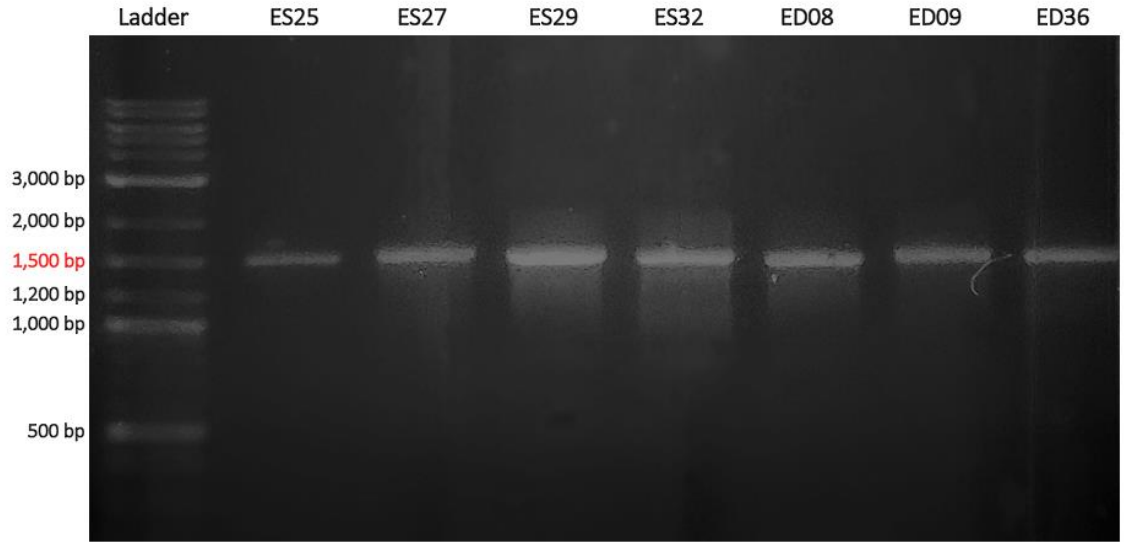

Figure 11. Gel electrophoresis of 16S E. faecalis PCR products from harbor seal and dog. Base pair length indicated next to the DNA ladder in well one. Each PCR product had a band size approximately 1,500 bp.

16S rRNA Phylogenetic Analysis

Isolates ES13, ES25, ES27, ES29, ES30, ES32, ED08, ED09, and ED36 each had a final sequence length of 1,433 base pair (bp) following manual curation in BioEdit. *E. faecalis* 16S sequences generated from harbor seal and dog scat had a 100% identity match, showing no differences in nucleotides between *E. faecalis* from these two sources. Each *E. faecalis* consensus sequence put into BLAST had the same top 100 accession

matches with percent identity scores ranging from 100-99.93%. The top matches were not exclusive to mammals and included sequences isolated from insects, fish, and soil. Phylogenetic analysis of *E. faecalis* isolates showed that isolates from different mammalian and non-mammalian sources clustered together, with each branch having a length of 0.00; this confirms that 16S rRNA sequencing does not provide the resolution needed to differentiate sources of *E. faecalis* contamination (Figure 12).

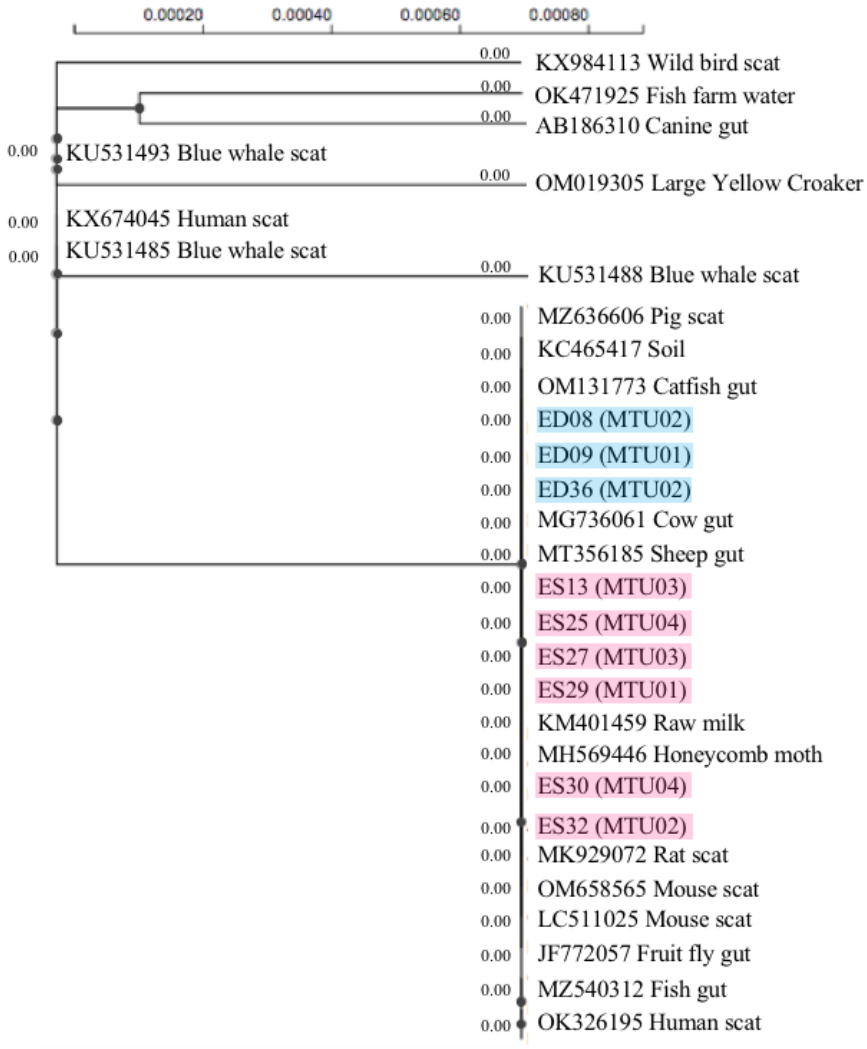

Figure 12. 16S phylogenetic analysis of E. faecalis isolated from harbor seal scat (pink) and domestic dog scat (blue).

CHAPTER IV:

DISCUSSION

MALDI-TOF MS

The objective of this study was to assess the capability of MALDI-TOF MS to differentiate Enterococci isolated from sources of environmental fecal pollution. 16S rRNA phylogenetic analysis, a gold standard for bacterial identification, was performed to confirm the identification of *E. faecalis* isolates identified through MALDI-TOF MS.

Distinct clusters were seen following cluster analysis of all isolates identified as Enterococci from harbor seal scat, domestic dog scat, and sewage (Figure 7). These clusters were consistent with the MTUs defined; of the 17 MTUs identified, only two MTUs (MTU03 and MTU05) had Enterococci isolated from different sources (Table 9). Cluster analysis of isolates reliably identified as *E. faecalis* from harbor seal and domestic dog scat resulted in four MTUs, two of which were harbor seal-specific (Figure 9). *E. faecalis* protein spectra from each MTU showed distinct peak differences (Appendix). MTUs defined through cluster analysis of all Enterococci isolated were mostly consistent with the MTUs defined for *E. faecalis* isolates. The harbor seal *E. faecalis* isolates in seal-specific MTUs were also in the seal-specific MTU (MTU15) defined by analysis of all isolates. This suggests that MALDI-TOF can reliably differentiate sources of environmental Enterococci pollution.

A similar study performed by Mirani (2022) assessed the capability of MALDI-TOF MS to differentiate sources of environmental *E. coli* pollution. This study found that MALDI-TOF MS could differentiate sources of *E. coli* isolated from human-composite from *E. coli* isolated from animal samples (harbor seal and domestic dog scat); however, distinct clusters could not be seen in *E. coli* isolated from harbor seal and dog scat, respectively. This may suggest that MALDI-TOF is more reliable in discriminating

sources of Enterococci pollution than *E. coli*; though, the lack of source discrimination between harbor seal and dog scat as shown by Mirani (2022) may be due to the low number of *E. coli* isolated from harbor seal scat with only four harbor seal-isolated *E. coli* being used.

MALDI-TOF MS and 16S rRNA Sequencing

MTUs defined through cluster analysis of Enterococci mass spectra suggests that MALDI-TOF MS may be a reliable method in multiple source tracking to identify sources of environmental Enterococci contamination. A library of 22 *E. faecalis* isolates clustered into four separate MTUs, with two MTUs being harbor seal-specific. The species identification given by the Bruker MALDI Biotyper system was confirmed as *E. faecalis* through 16S rRNA sequencing. Furthermore, phylogenetic analysis of *E. faecalis* isolated from different environmental sources confirms that 16S rRNA sequencing does not provide the resolution needed for source discrimination of environmental *E. faecalis* isolates.

16S rRNA Phylogenetic Analysis

16S rRNA sequencing of isolates reliably identified as *E. faecalis* by the Bruker Biotyper system verified that MALDI-TOF MS can accurately identify these isolates. Phylogenetic analysis of *E. faecalis* 16S rRNA sequences from harbor seal and dog scat showed that sequences from different mammalian and non-mammalian sources clustered together, confirming that 16S rRNA sequencing is unable to differentiate sources of *E. faecalis* contamination. Following manual curation, the DNA sequences generated from 16S rRNA sequencing had a final length of 1,433 bp; this is close to the length of the full 16S gene (approximately 1,500 bp). Despite sequencing the near full 16S gene, the sequences generated from both harbor seal and domestic dog *E. faecalis* isolates were

identical. This confirms that 16S rRNA sequencing is unable to discriminate environmental sources of *E. faecalis* pollution.

Current Limitations

Sample Replication

A key limitation of this study includes the number of samples used to culture Enterococci. The library of Enterococci generated in this study was isolated from a single composite sample of harbor seal scat, a sample of domestic dog scat, and a composite sample of human waste from a waste treatment facility. Of these single samples, only one gram was used to create the sample slurries used for serial dilutions. This sample limitation provides a very small window into the overall Enterococci composition of these samples; this may have also contributed to the low species diversity of Enterococci isolated from each source, particularly harbor seal scat and sewage samples (two and three species isolates, respectively). Generating a larger library of Enterococci isolated from more samples of the sources used would provide more species-specific mass spectra to use for comparison between sources, such as was done in this study with *E. faecalis* isolated from harbor seal and dog scat.

Though the preliminary 2018 study suggested that gut bacteria from captive harbor seals are closely related enough to use for studying wild seals, this study would be improved by using wild seal scat to isolate Enterococci for MALDI-TOF MS analysis.

Reference Databases

A limitation in the application of MALDI-TOF MS for MST is amount of environmental protein mass spectra in reference databases. To accurately identify microorganisms using protein mass spectra, a reference spectra of the same strain is needed. The Bruker MALDI Biotyper database is diverse; however, it primarily consists of clinically-isolated microorganisms. The addition of more environmental sources of

Enterococci mass spectra would provide better discrimination between sources as there would be more spectra peaks in the reference databases to compare to.

Future Directions

The issue of beach closures in Cape Cod, MA primarily persists due to the lack of contamination source identification. Having the ability to find where fecal indicators originate from would allow officials to identify the reason for exceedances in fecal contamination. Pinpointing the source of fecal contamination and its associated transmission pathway into recreational water would allow beach administrators to identify actions needed for remediation.

MALDI-TOF MS could be implemented to further expand upon the EPA's current water quality methodology. Though the formic acid/ethanol tube extraction method helps ensure protein is extracted for analysis, MALDI-TOF MS can be performed with isolates taken directly from a fresh agar plate. With an in-house MALDI-TOF MS system, plating directly from an agar plate onto the steel target and performing MALDI-TOF MS would provide researchers with mass spectra data for Enterococci isolates within an hour. With the low cost of MALDI-TOF reagents per protein extraction and the short turnaround time for mass spectra generation, MALDI-TOF MS could be used to build a library of mass spectra from environmentally-isolated fecal indicators. Identifying peaks in mass spectra specific to Enterococci based on isolation source would allow researchers to easily identify the potential source of fecal contamination.

REFERENCES

Ahmed W, Stewart J, Gardner T, Powell D (2008) A real-time polymerase chain reaction assay for quantitative detection of the human-specific enterococci surface protein marker in sewage and environmental waters. Environ Microbiol. 10:3255–3264.

Arnold BF, Wade TJ, Benjamin-Chung J, et al. (2016) Acute Gastroenteritis and Recreational water: highest burden among young US children. Am J Public Health. 106:1690-1697.

AVMA (2018) AVMA Pet Ownership and Demographics Sourcebook*:* 2017-2018 Edition. Schaumburg, IL: American Veterinary Medical Association.

Barberis C, Almuzara M, Join-Lambert O, Ramírez MS, Famiglietti A, Vay C (2014) Comparison of the Bruker MALDI-TOF mass spectrometry system and conventional phenotypic methods for identification of gram-positive rods. PLOS ONE 9:e106303.10.1371/journal.pone.0106303

Bernhard AE, Field KG (2000) Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Appl. Environ. Microbiol. 66:1587–1594.

Byappanahalli MN, Nevers MB, Korajkic A, Staley ZR, Harwood VJ (2012) Enterococci in the environment. Microbiol Mol Biol Rev. 76:685-706.

Costalago D, Bauer B, Tomczak MT, et al. (2019) The necessity of a holistic approach when managing marine mammal–fisheries interactions: Environment and fisheries impact are stronger than seal predation. Ambio 48:552–564.

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Research.

36:W465–W469.

Drost (2018) Philentropy: Information theory and distance quantification with R. Journal of Open Source Software. 3:765.

Elmir SM, Wright ME, Abdelzaher A, Solo-Gabriele HM, Fleming LE, Miller G, et al. (2007) Quantitative evaluation of bacteria released by bathers in a marine water. Water Res. 41:3–10.

EPA (1998) How Wastewater Treatment Works ... The Basics. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water. 833-F-98-002.

Fiksdal L, Maki JS, LaCroix SJ, Staley JT (1985) Survival and detection of *Bacteroides* spp., prospective indicator bacteria. Appl Environ Microbiol. 49:148–150.

Frungillo, J (2014) An analysis of gray and harbor seal strandings in Cape Cod, Massachusetts from 1999 to 2012.

Gibb S, Strimmer K (2012) MALDIquant: a versatile R package for the analysis of mass spectrometry data. Bioinformatics. 28:2270–2271.

Giebel RA, Fredenberg W, Sandrin TR (2008) Characterization of environmental isolates of *Enterococcus* spp. By matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Water Res. 42:931-940.

Glad T, Fam Kristiansen V, Nielsen KM, Brusetti L, Wright A-DG, Sundset MA (2010) Ecological characterisation of the colonic microbiota in Arctic and sub-Arctic seals. Microbiol Ecol. 60:320–330.

Gómez-Doñate M, Casanovas-Massana A, Muniesa M, Blanch AR (2016) Development of new host-specific Bacteroides qPCRs for the identification of fecal contamination sources in water. Microbiologyopen 5:83-94.

Gorman M (1999) Environmental Hazards: Marine Pollution. Boulder, CO: NetLibrary, Inc.

Griffin PM, Price GR, Schooneveldt JM, Schlebusch S, Tilse MH, Urbanski T, Hamilton B, Venter D (2012) Use of matrix-assisted laser desorption ionization-time of flight mass spectrometry to identify vancomycin-resistant enterococci and investigate the epidemiology of an outbreak. Journal of Clinical Microbiology 50:2918-2931.

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

Halliday E, Gast RJ (2011) Bacteria in beach sands: An emerging challenge in protecting coastal water quality and bather health. Environ Sci Technol. 45:370–379.

Hettick JM, Kashon ML, Slaven JE, et al (2006) Discrimination of intact mycobacteria at the strain level: a combined MALDI-TOF MS and biostatistical analysis. Proteomics. 6:6416-6425.

Hsieh TC, Ma KH, Chao A (2016) iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods Ecol Evol. 7:1451-1456.

Huber F, Ridder L, Verhoeven S, Spaaks JH, Diblen F, et al. (2021) Spec2Vec: Improved mass spectral similarity scoring through learning of structural relationships. PLOS Computational Biology. 17:e1008724.

Hudak CA, Sette L (2019) Opportunistic detection of anthropogenic micro debris in harbor seal (Phoca vitulina vitulina) and gray seal (Halichoerus grypus atlantica) fecal samples from haul-outs in southeastern Massachusetts, USA. Marine Pollution Bulletin. 145:390-395.

Janda JM, Abbott SL (2007) 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J Clin Microbiol. 45:2761-

2764.

Jin G, Englande A, Bradford H, and Jeng HW (2004) Comparison of E.Coli, Enterococci, and fecal coliform as indicators for brackish water quality assessment. Water Environment Research. 76:245-255.

Johnson JS, Spakowicz DJ, Hong BY et al. (2019) Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun 10**:**5029.

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. doi:10.7717/peerj.11359

Layton A, Mckay L, Williams D, Garrett V, Gentry R, Sayler G (2006) Development of Bacteroides 16S rRNA Gene TaqMan-Based Real-Time PCR assays for estimation of total, human, and bovine fecal pollution in water. Appl. Environ. Microbiol. 72:4214–4224.

Lelli B, Harris DE, Aboueissa AM (2009) Seal bounties in Maine and Massachusetts, 1888 to 1962. Northeastern Naturalist. 16:239–254.

Lisle JT, Smith JJ, Edwards DD, McFeters GA (2004) Occurrence of microbial indicators and *Clostridium perfringens* in wastewater, water column samples, sediments, drinking water, and Weddell seal feces collected at McMurdo Station, Antarctica. Appl Environ Microbiol. 70:7269-7276.

Mazhari F (2021) Application of whole genome sequencing and MALDI-TOF to identification of Bacillus species isolated from cleanrooms at NASA Johnson Space Center.

Medeiros AW, Blaese Amorim D, Tavares M, et al. (2017) *Enterococcus* species

diversity in fecal samples of wild marine species as determined by real-time PCR. Can J Microbiol. 63:129-136.

Mirani A (2022) Application of MALDI-TOF MS for multiple source tracking from sewage and septic tanks along with identification of antibiotic resistant pathogenic *Escherichia coli*.

Moxley J, Bogomolni A, Hammill M, Moore K, Polito M, Sette L, Johnston D, et al. (2017) Google haul uut: Earth observation imagery and digital aerial surveys in coastal wildlife management and abundance estimation. BioScience. 67:760–768.

Nelson TM, Rogers TL, Carlini AR, Brown MV (2013) Diet and phylogeny shape the gut microbiota of Antarctic seals: a comparison of wild and captive animals. Environ Microbiol. 15:1132–1145.

Numberger D, Herlemann DPR, Jurgens K, Dehnhardt G, Schulz-Vogt H (2016) Comparative analysis of the fecal bacterial community of five harbor seals (Phoca vitulina). Microbiologyopen. 5:782–792.

Orskov F, Orskov I (1981) Enterobacteriaceae, p. 340-352. In A. I. Broude (ed.), Medical microbiology and infectious diseases. W. B. Saunders Co., Philadelphia, Pa.

Pace RM, Josephson E, Wood SA, Murray K, Waring G (2019) Trends and patterns of seal abundance at haul-out sites in a gray seal recolonization zone. NOAA technical memorandum NMFS-NE; 251.

Pacheco-Sandoval A, Schramm Y, Heckel G, Brassea-Pérez E, Martínez-Porchas M, Lago-Lestón A (2019) The Pacific harbor seal gut microbiota in Mexico: Its relationship with diet and functional inferences. Plos One. 14. doi:10.1371/journal.pone.0221770

Panda A, Kurapati S, Samantaray JC, Srinivasan A, Khalil S (2014) MALDI-TOF mass

spectrometry proteomic based identification of clinical bacterial isolates. Indian J Med Res. 140:770-777.

Reeves RL, Grant SB, Mrse RD, Copil Oancea CM, Boehm AB (2004) Sources, management, and scaling of fecal indicator bacteria in runoff from a coastal urban watershed in Southern California. Environ Sci Technol. 38:2637–2648.

Sauer S, Freiwald A, Maier T, Kube M, Reinhardt R, Kostrzewa M, Geider K (2008) Classification and identification of bacteria by mass spectrometry and computational analysis. PLOS ONE 3:e2843.10.1371/journal.pone.0002843

Scott TM, Rose JB, Jenkins TM, Farrah SR, Lukasik J (2002) Microbial Source Tracking: Current Methodology and Future Directions. Appl. Environ. Microbiol. 68:5796-5803.

Seng P, Rolain JM, Fournier PE, La Scola B, Drancourt M, Raoult D (2010) MALDI-TOF-mass spectrometry applications in clinical microbiology, Future Microbiology. 5:1733+.

Sette L, Accardo CM, McKenna K, Patchett K, Rose K, Sharp BW, Waring GT (2020) The seasonal diet, distribution, and counts of harbor seals (Phoca vitulina vitulina) and gray seals (Halichoerus grypus atlantica) in Pleasant Bay and Chatham Harbor, Cape Cod, Massachusetts. Northeastern Naturalist. 27:98–113.

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.

Singhal N, Kumar M, Kanaujia PK, Virdi JS (2015) MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. Frontiers in

Microbiology 6:e791.10.3389/fmicb.2015.00791

Smith SC, Chalker A, Dewar ML, Arnould JPY (2013) Age-related differences revealed in Australian fur seal *Arctocephalus pusillus* doriferus gut microbiota. FEMS Microbiol Ecol. 86:246–255.

Staley ZR, Boyd RJ, Shum P, Edge TA (2018) Microbial source tracking using quantitative and digital PCR to identify sources of fecal contamination in stormwater, river water, and beach water in a great lakes area of concern. Appl Environ Microbiol. 84:e01634-18.

Steele JA, Blackwood AD, Griffith JF, Noble RT, Schiff KC (2018) Quantification of pathogens and markers of fecal contamination during storm events along popular surfing beaches in San Diego, California. Water Res. 135:137-149.

Suzuki R, Shimodaira H (2006) Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics. 22:1540–1542.

Vaz-Moreira I, Egas C, Nunes OC, Manaia CM (2011) Culture-dependent and culture-independent diversity surveys target different bacteria: a case study in a freshwater sample. Antonie Van Leeuwenhoek. 100:245-257.

Wang CC, Lai YH, Ou YM, Chang HT, Wang YS (2016) Critical factors determining the quantification capability of matrix-assisted laser desorption/ionization- time-of-flight mass spectrometry. Philos Trans A Math Phys Eng Sci. 374:20150371.

Wexler AG, Goodman AL (2017) An insiders perspective: Bacteroides as a window into the microbiome. Nature Microbiology. doi:10.1038/nmicrobiol.2017.26.

Wolters M, Rohde H, Maier T, et al (2011) MALDI-TOF MS fingerprinting allows for discrimination of major methicillin-resistant *Staphylococcus aureus* lineages. Int J Med Microbiol. 301:64-68.

Wright ME, Solo-Gabriele HM, Elmir S, Fleming LE (2009) Microbial load from animal feces at a recreational beach. Mar Pollut Bull. 58:1649-1656.