Comparing Two Non-Invasive Methods for Assessing Marine Mammal Genetic Diversity: Environmental DNA vs. Fecal DNA

Sydney Jackson
COMPARING TWO NON-INVASIVE METHODS FOR ASSESSING MARINE MAMMAL GENETIC DIVERSITY: ENVIRONMENTAL DNA VS. FECAL DNA

by

Sydney Jackson

A Thesis Submitted in Partial Fulfillment of the Requirements for a Degree with Honors (Marine Science)

The Honors College
University of Maine
April 2022

Advisory Committee:
Kristina Cammen, Assistant Professor of Marine Mammal Science
James Brophy, Preceptor, Honors College
Sindhu Manjesh, Lecturer, Journalism
Julia Sunnarborg, PhD Student, Marine Biology
ABSTRACT

As technology and science progresses, the methodology behind observing, monitoring, and sampling marine mammals advances as well. One such technique is environmental DNA or eDNA, which entails extracting organismal DNA from water samples without ever handling or disturbing the organism. It is a cost-efficient and non-invasive method that can be utilized in the sampling of seal haulout sites as is its purpose for this research. Another method, using the DNA analysis of seal fecal samples, is a less invasive method that can also be utilized to monitor and assess marine mammals. Through collecting both fecal and water samples from gray seal haulout sites in Cape Cod, Massachusetts, these two differing, but equally progressive methods can be compared to one another. The water samples collected from the seal haulout sites were paired for DNA analysis with the fecal samples collected from the beaches where gray seals are hauled out in Cape Cod. DNA was then extracted from both the water samples and fecal samples, followed by sequencing a portion of the gray seal mitochondrial control region in all the samples. This allowed for the comparison of the haplotypes detected in fecal samples to those detected in water samples as a comparison of these two non-invasive approaches for assessing marine mammal genetic diversity. We obtained sequences from 25 fecal samples. Sequences from all but 2 of the 25 samples were found to match with one of the sequences in the reference dataset. Our study identified 2 new haplotypes that had not been previously identified in the population. When compared to the water sample sequences, we found 19 matches out of the 25 fecal sample sequences. In all of these cases, the fecal haplotype was detected in water samples collected during
the same survey (at the same haulout on the same day), though in many cases a given fecal haplotype was also detected in water samples from multiple surveys. Although future studies are needed to further confirm the efficiency and non-invasiveness of the eDNA approach, our study suggests that it can provide similar information to a fecal sample sequence analysis, but in a less invasive way.
ACKNOWLEDGEMENTS

I would like to thank Christy Hudak and Lisa Sette from the Center for Coastal Studies for providing the samples and data used in this thesis. I would also like to thank my thesis advisor, Kristina Cammen, and Julia Sunnarborg for supporting and encouraging me throughout the entirety of this process as well as my committee members, Walter Golet, Michael Kinnison, James Brophy, and Sindhu Manhjesh, for all their feedback, advice, and support they provided for me during this thesis process. Finally, I want to thank all my friends and family for being an amazing support system for me and for always giving me words of encouragement when I needed it.

Funding was provided through the UMaine Center for Undergraduate Research in the form of a grant.
# TABLE OF CONTENTS

INTRODUCTION.............................................................................................................. 1

Introduction to eDNA................................................................................................. 1

The study species: gray seals.................................................................................... 3

Study Objectives........................................................................................................... 4

METHODS....................................................................................................................... 6

Sample Collection....................................................................................................... 6

Sample Processing..................................................................................................... 7

Sequencing................................................................................................................... 8

Data Analysis.............................................................................................................. 9

RESULTS......................................................................................................................... 12

Fecal DNA vs. Reference Database............................................................................ 12

Fecal DNA vs. eDNA.................................................................................................. 12

DISCUSSION.................................................................................................................... 14

REFERENCES ................................................................................................................ 18

APPENDICES ................................................................................................................ 20

APPENDIX A: FIGURES................................................................................................ 21

APPENDIX B: TABLES................................................................................................... 23

AUTHOR’S BIOGRAPHY............................................................................................... 25
LIST OF FIGURES

Figure 1: Map of gray seal haulout sites in Cape Cod, MA. 21
Figure 2: Methods flowchart 22

LIST OF TABLES

Table 1: Appearance of water sample sequences in each survey 23
Table 2: Comparison of number of water and fecal sequences in each survey. 24
INTRODUCTION

Introduction to eDNA

Environmental DNA, also known as eDNA, is organismal DNA that can be found in aquatic or terrestrial environments. Organisms leave a trace of DNA in the environments they inhabit through lost skin cells, bodily fluids, excrement, and other forms of DNA. These remnants can be sampled and analyzed using new molecular methods that are highly sensitive to low quantities of DNA.

eDNA is able to be utilized for a variety of different research questions, including, but not limited to, single-species detection, identification and protection of rare or protected species, assessing genetic diversity in common species, and characterizing community composition. One study, conducted by Foote et al. (2012), investigated the potential use of eDNA for the genetic monitoring of marine mammals. These researchers utilized specific primers to amplify short mitochondrial DNA sequences to attempt to detect the presence of the harbor porpoise, *Phocoena phocoena*, as well as compare these detections to those of harbor porpoise echolocation clicks. The results of this study indicated that although detection by eDNA was less successful than acoustic detections, eDNA has the potential to be just as successful as current visual and acoustic methods of species detection of marine mammals with the proper optimization of larger volumes of seawater (Foote et al., 2012). A second study, conducted by Sigsgaard et al. (2017), aimed to demonstrate that high-throughput sequencing of seawater eDNA can be used to approximate the genetic diversity of whale shark (*Rhincodon typus*) aggregations. They found that there were similar mitochondrial haplotype frequencies in seawater compared...
to tissue samples, thus further validating the role eDNA plays in assessing population genetics of aquatic organisms (Sigsgaard et al. 2017). Berry et al. (2019) investigated the use of eDNA to measure changes in the biological composition of communities in different ocean regions, specifically in zooplankton species. Findings from this study included the identification of consistent seasonal assemblages of zooplankton species, thus indicating the efficiency of the eDNA method in surveying community composition (Berry et al., 2019). These studies demonstrate the wide scope of research questions that eDNA is capable of answering, ranging from species detection to understanding changes in community composition.

eDNA provides several advantages to scientific researchers, including its non-invasiveness, potential for citizen science, ability to sample whole communities, as well as its cost-effectiveness. Methods designed to genetically characterize an organism prior to eDNA included capturing species and taking skin, blood, and other invasive samples. This not only involves having to disturb an organism’s environment, but disturbs and potentially harms the organism themselves. In some circumstances, even lethal sampling has been used. eDNA, on the other hand, allows for researchers to collect data with minimal disturbance of the organism or its environment. When a water sample is collected, organismal DNA can be extracted from it without ever handling the organism. These methods appeal especially to marine mammal specialists, whose goal is to be able to sample populations in a non-invasive way utilizing eDNA.

The advantages that eDNA provide are significant, particularly for studying and understanding a variety of marine mammal species. One study, conducted by Andruszkiewicz et al. (2017), utilized the eDNA metabarcoding method to identify
vertebrate communities at multiple oceanographic stations within the Monterey Bay National Marine Sanctuary. This study found that all but 1 family identified using eDNA metabarcoding is known to exist in the Monterey Bay National Marine Sanctuary, thus confirming the accuracy and importance of eDNA metabarcoding for vertebrate biomonitoring (Andruszkiewicz et al., 2017). A second study, done by Baker et al. (2018), utilized droplet digital PCR technology using eDNA collected from seawater for detection and species identification of cetaceans. The findings of this study confirmed the possibility to detect eDNA in the wake of whales, thus further validating the efficiency of the eDNA process in assessing marine mammal populations (Baker et al., 2018). Another study conducted by Szekely et al. (2021) is one of the most recent studies done regarding the use of eDNA for marine mammal studies. In this study, researchers investigated whether eDNA isolated from seawater samples was able to be used to assess the genetic diversity of bowhead whales in West Greenland. The findings of this study indicate that utilizing the eDNA method to analyze samples collected in the footprint or wake of migrating animals has potential to accurately assess the genetic diversity of bowhead whales and other marine mammals (Szekely et al., 2021). Despite the successes of these studies, one gap in the literature surrounding marine mammals is research regarding pinnipeds and how beneficial the eDNA method could be in helping to understand their populations better. One particular pinniped population of interest is the gray seal (Halichoerus grypus) population.

**The study species: gray seals**

The gray seal (*Halichoerus grypus*) is a prevalent pinniped in New England and specifically, Cape Cod, Massachusetts. Gray seals are the largest seal found in the Cape
Cod area, with males growing to 8 feet and weighing over 800 pounds and females measuring approximately 7 feet and weighing less than 600 pounds. Males also tend to be darker with few light spots while females tend to be light with dark, irregular blotches. Gray seals are found primarily in the North Atlantic, ranging from the Baltic, western Europe to Canada and Northeastern United States (Katona et al., 1993)

Gray seals are very social pinnipeds that accumulate on shared terrestrial haulouts in between single- and multi-day movements offshore (Moxley et al., 2020). These haulouts provide an environment for scientists to collect fecal samples from gray seals, as well as water samples off their coasts, as was done in this study. Similar to eDNA analyses, researchers can recover low quantities of seal DNA from a fecal sample that includes sloughed intestinal wall skin cells from the animal. The analysis of DNA from these samples can help researchers to learn more about the gray seals on that haulout, including measures of population genetic diversity.

The gray seal species was declared protected after the passage of the U.S. Marine Mammal Protection Act of 1972. Since then, the gray seal population in the Northwest Atlantic has grown exponentially, recovering from a historical bottleneck due to human exploitation and bounties (Wood et al., 2020). A population bottleneck is a drastic reduction in the size of a population that can have effects on fitness and potential for future adaptation in natural populations. The analysis of genetics allows for a deeper understanding surrounding the loss of genetic diversity during one of these events, as well as the subsequent recovery of the species affected (Cammen et al., 2018).

**Study Objectives**
This study aims to compare two minimally invasive approaches of assessing marine mammal genetic diversity. Specifically, we will compare the genetic sequences detected in gray seal fecal samples to those detected in eDNA water samples. eDNA is a cost-efficient and non-invasive method that can be utilized in both terrestrial and aquatic environments, but appeals especially to marine mammal scientists who are constantly trying to develop less invasive methods of collecting and analyzing data. Fecal DNA analysis is another minimally invasive method that can be used to assess and monitor marine mammal populations. While eDNA sampling can be completed from a boat at a distance from the haulout, fecal DNA collection requires collecting samples from the haulout itself, which results in temporarily displacing the animals. Boat approaches for eDNA collection can also result in temporary displacement if the seals are disturbed by boat presence, but generally the eDNA water sample collection process is less disruptive.

Through the comparison of these two minimally invasive sampling methods, fecal DNA and eDNA, this study aims to contribute to the already growing knowledge surrounding eDNA and other less invasive approaches to assessing marine mammal genetic diversity. As fecal DNA analysis has been previously demonstrated to be successful across many study species, this comparison is useful to validate the effectiveness of the more novel eDNA approach. The hope of this research is to gain enough evidence to further support the efficient use of eDNA for future marine mammal genetic studies.
METHODS

To complete these objectives, I played a role in a larger collaborative research project. Collection and filtering was done by the Center for Coastal Studies. Students in the Cammen Lab completed fecal and water sample extractions and conducted qPCR analyses. Once samples were sequenced, I processed the fecal sample sequences and analyzed the data.

Sample Collection

Seawater and seal fecal samples were collected at several haulout sites in Cape Cod, Massachusetts between June and September, 2020 by collaborators at the Center for Coastal Studies (Fig. 1). The haulouts were approached by boat, and a transect was followed parallel to the shore at a distance of 50 m for visual monitoring of all animals prior to approaching the haulout for sampling on shore. Water samples were collected 50 meters off the coast of the gray seal haul outs at the start and end of the transect, as well as from the beach at the haulout mid point after the gray seals had flushed into the water following the boat’s approach. At each point, two 1 liter bottle samples of surface seawater were collected, and stored on ice in a cooler. All fecal samples on the beach were then collected by hand and stored individually in a separate cooler from the water samples.

In total, 10 sampling surveys were conducted in 2020, during which both fecal and water samples were collected. However, water samples collected during Survey 1 could not be sequenced due to a contamination issue that occurred in the lab during the
DNA extraction of these samples. We therefore include fecal samples from all 10 surveys, but consider only 9 surveys with paired fecal and water sampling in our study.

**Figure 1**: Map of Cape Cod, Massachusetts displaying the location of the 6 gray seal haulout sites where samples were collected.

**Sample Processing**

Within the lab, each water sample was filtered through a cellulose nitrate filter with a pore size of 0.45 m to capture free-floating DNA molecules (Fig. 2). Lab blanks of 1 L of tap water were also filtered both at the beginning and the end of the filtering process. These blanks serve as negative controls to test for contamination during the filtration process. All filtering took place on a vacuum powered manifold.
DNA was extracted from the filters using a Qiagen DNeasy blood and tissue extraction kit and from the fecal samples using a QIAamp DNA stool mini kit (Fig. 2). Extraction blanks were included in all extractions done. This is to detect any contamination that might have been introduced during the extraction process. All DNA extractions were then stored at -20 degrees C.

Following extraction, quantitative PCR with seal-specific probes was used to determine which water samples contained seal eDNA. All positive water samples along with the fecal samples were then prepared for sequencing.

**Sequencing**

Once samples are collected, researchers use a variety of sequencing methods to access the specific genetic data they need to investigate their research questions. This study uses both Sanger sequencing and next-generation sequencing. Sanger sequencing is appropriate when the genetic material in a sample comes from a single individual (i.e., a pure sample), while next-generation sequencing is useful when a sample may contain a mix of genetic material from multiple individuals (i.e., a mixed sample). Accordingly, in our study, fecal samples are analyzed using Sanger sequencing and water eDNA samples are analyzed using next-generation sequencing (Fig. 2). Both methods require that the targeted DNA is first amplified to achieve a higher concentration using primers that target a specific region of the genome, in a process known as polymerase chain reaction (PCR). The sequencing approaches then determine the order of nucleotides (i.e., DNA building blocks) in single-stranded DNA molecules, with Sanger sequencing producing a single sequence as a result and next-generation sequencing producing many thousands of sequence reads that are analyzed bioinformatically (Fig. 2).
In this study, we sequenced a 423 base pair fragment of the mitochondrial control region in order to identify gray seal haplotypes present in these samples. Haplotypes are a set of DNA variations along a chromosome that tend to be inherited together because they're very close together. Previous research has identified 30 unique mitochondrial control region haplotypes in the gray seal population in the Northwest Atlantic (Cammen et al. 2018; Wood et al. 2011).

![Flowchart](image)

**Figure 2:** Flowchart beginning with sample collection, followed by filtering, DNA extraction, and sequencing.

**Data Analysis**

Once the DNA sequences were returned, the forward and reverse sequence reads from each fecal sample were paired and then visually inspected using Codon Code
Aligner in order to remove primers and correct mismatches and errors. This process resulted in a single inferred haplotype for each fecal sample.

Fecal haplotypes were compared to published sequences from Cammen et al. (2018), and to sequences from the water samples collected at the same sites in Cape Cod. The sequences from Cammen et al. (2018) represent haplotypes that were previously identified in tissue samples collected from gray seals from Massachusetts to Canada. We consider these a reference dataset because it was derived from tissue samples, which are the standard used to characterize genetic diversity of an individual. There are 38 distinct haplotypes present in this reference dataset, which were derived from a total of 385 gray seal individuals included in the prior study. One caveat to note in this comparison is that the region sequenced differed slightly between our study and this prior study. Because the studies used different primers, when the sequences were aligned, our sequences did not cover the first 18 basepairs of the haplotype sequences in the reference dataset and our sequences included an additional 17 basepairs at the end that were not included in the reference dataset. It is therefore possible for our sequences to match multiple reference haplotypes, if the reference haplotypes differ in the first 18 basepairs not included in our sequence.

The comparison with haplotypes from the water eDNA samples collected concurrently with the sequenced fecal samples was more straightforward, as both fecal samples and water eDNA samples were sequenced with the same primers. Using the paired nature of collecting fecal and water samples from the same haulout site on the same day, we determined whether or not the water sample sequences correlating with the fecal sample sequences were observed in the same survey number as the fecal sample
sequences, as well as the total number of surveys the water sample sequences were observed in.
RESULTS

We obtained sequences from 25 fecal samples. We first compared these sequences to one another. The samples produced a total of 17 unique haplotypes, with 8 haplotypes observed in two samples each.

Fecal DNA vs. Reference Database

We then compared the fecal sequences to the reference database. Sequences from all but 2 of the 25 samples were found to match with one of the sequences already published in the Cammen et al. (2018) reference dataset. Our study therefore identified 2 new haplotypes that had not been previously identified in the population.

Fecal DNA vs. eDNA

Finally, we compared the fecal sequences to the water eDNA sequences. When compared to the water sample sequences, we found 19 matches out of the 25 fecal sample sequences. In all of these cases, the fecal haplotype was detected in water samples collected during the same survey (at the same haulout, on the same day) (Table 2) though in many cases a given fecal haplotype was also detected in water samples from multiple surveys (Table 1). 11 fecal haplotypes were found in 9 out of the 9 surveys, 7 were found in 8 of the 9 surveys, and 1 was found in 4 out of the 9 surveys.

In total, there were 6 fecal sample sequences that did not match any of the water sample sequences, indicating that the genetic material from these individuals on the beach were not detected in the nearby water collections.
Table 1: Appearance of fecal sample sequences, as shown by matching ASV, in each survey, indicated by Y(Yes) and N(No). This Y/N was based off of any reads in any of the replicates (some had only very few reads). Survey 1 is blocked out because no water samples from this survey were sequenced.

<table>
<thead>
<tr>
<th>Survey</th>
<th>Survey 2</th>
<th>Survey 3</th>
<th>Survey 4</th>
<th>Survey 5</th>
<th>Survey 6</th>
<th>Survey 7</th>
<th>Survey 8</th>
<th>Survey 9</th>
<th>Survey 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>3</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>4</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>6</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>7</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>8</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>9</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>10</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

Table 2: Comparison of the number of fecal samples and number of matching eDNA sequences observed in each survey. N/A under # Fecal sequences observed indicates that no fecal samples were collected in that survey, while N/A under # eDNA sequences observed in the same survey as matching fecal sequences indicates that there were no matching eDNA sequences identified.

<table>
<thead>
<tr>
<th>Survey</th>
<th>Haulout Site</th>
<th>Date Collected</th>
<th># Fecal sequences observed</th>
<th># eDNA sequences observed in same survey as matching fecal sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CH</td>
<td>6/23/2020</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>PT</td>
<td>6/25/2020</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>CH</td>
<td>7/6/2020</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>PB</td>
<td>7/6/2020</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>PH</td>
<td>7/21/2020</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>TI</td>
<td>8/3/2020</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>SI/SI-PLB</td>
<td>8/3/2020</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>SI</td>
<td>8/19/2020</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>9</td>
<td>LP</td>
<td>8/25/2020</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>TI</td>
<td>9/17/2020</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>
DISCUSSION

The main objective of this study was to compare two minimally invasive approaches of assessing marine mammal genetic diversity through the analysis of genetic sequences detected in both gray seal fecal samples as well as in the water surrounding gray seal haulout sites in Cape Cod, Massachusetts. Other methods of studying marine mammals prior to eDNA were typically very invasive and disruptive to the species being studied, usually involving capturing species and collecting blood, skin, and other invasive samples. eDNA-based approaches, which were only recently proposed as an alternative, have so far shown some promise in characterizing the genetic diversity of cetaceans when water samples are collected in their close vicinity. A study conducted by Sigsgaard et al. (2017) aimed to demonstrate that high-throughput sequencing of seawater eDNA can be used to approximate the genetic diversity of whale shark (Rhincodon typus) aggregations. They found that there were similar mitochondrial haplotype frequencies in seawater compared to tissue samples, thus further validating the role eDNA plays in assessing population genetics of aquatic organisms (Sigsgaard et al. 2017).

The application of eDNA in pinnipeds has been lacking thus far, and there is no published comparison of eDNA-based approaches to fecal DNA sampling. Through this comparison, we hoped to demonstrate that eDNA-based approaches are just as effective in characterizing gray seal genetic diversity as fecal sample DNA analysis. If this proved to be true, the less invasive process of eDNA could potentially be utilized even more effectively within the marine mammal scientific community as well as in a variety of other scientific studies.
When analyzing the overlap between fecal sample and water sample sequences, we found a relatively high number of matches. 76% or 19 out of 25 fecal sample sequences matched sequences also identified in the water samples. Further analysis of how many water sample sequences were observed in the same survey as their respective fecal sample sequence match also revealed promising findings. Every matching water sample sequence was found in the same survey number as their respective fecal sample sequence, confirming that the matches occur when samples are collected in overlapping time and space (Table 2). This is important because we want to know that eDNA data from water samples reflects the genetic diversity of the seals at that beach at that time.

There is still a lot of uncertainty about how long an eDNA signal lasts in an environment. Székely et al. (2021) found that bowhead whale eDNA collected in a footprint of a diving whale is hard to detect after only 10 minutes, but Baker et al. (2018) detected killer whale eDNA up to 2 hours after a pod traveled through an area. More scientific studies are needed in this area to better understand eDNA persistence around seal haulouts if eDNA-based monitoring of marine mammal species is to be implemented further.

We also found that most of our fecal sequences were detected in multiple water samples, including those collected at different beaches on different days (Table 1). This likely reflects that multiple individuals can share the same haplotype. In our study, we detected up to two fecal samples with a shared haplotype. In a prior study of tissue samples, common haplotypes were identified in over 20 individuals from the same population (Cammen et al. 2018). This highlights that there is not necessarily a one-to-one relationship between an eDNA-derived sequence and a single individual seal, which
will need to be taken into consideration if implemented in future monitoring. When a fecal sample sequence matches a water sample sequence, this indicates that both of these approaches are able to detect similar types of marine mammal genetic diversity. However, because multiple individuals can share the same haplotype, we cannot know if the matching sequence in the water sample originated from the same individual that produced the fecal sample.

In contrast to the scenario of fecal sequences matching water samples, when there is no water sample sequence that matches with a fecal sample sequence, we can conclude that the eDNA sample missed a haplotype present on that beach. In those cases, the water did not capture enough haplotypes to provide the information one would get using fecal DNA analysis. One of the limitations to fecal DNA approaches are that they can only assess the genetic diversity of animals that leave behind a fecal sample. At our sites, only a few fecal samples (up to 8) were found on each beach, despite counting up to 325 seals on the haulouts when they were sampled in 2020. eDNA-based approaches may be able to capture the genetic diversity of a greater number of these individuals. Preliminary analyses suggest that the water samples capture a large number of haplotypes that we presume represent DNA from multiple individuals, including those that did not leave a fecal sample on the beach. However, because our study did not involve a complete analysis of the sequences derived from eDNA water samples, we cannot fully evaluate what fecal sampling misses in comparison to water sampling at this time.

Although future studies are needed to further confirm the efficiency and non-invasiveness of the eDNA approach, our study suggests that it can provide similar information as a fecal sample sequence analysis, but in a less invasive way. Prior to this
study, little research had been done in order to determine how beneficial the eDNA approach could be to understanding the pinniped population. Utilizing an approach that requires little to no disruption of the marine mammal species being studied while also gathering necessary information is an essential step in not only the marine mammal field of study, but more specifically, the pinniped population.
REFERENCES


Figure 1: Map of Cape Cod, Massachusetts displaying the location of the 6 gray seal haulout sites where samples were collected.
Figure 2: Flowchart beginning with sample collection, followed by filtering, DNA extraction, and sequencing.
APPENDIX B: TABLES

<table>
<thead>
<tr>
<th>ASV</th>
<th>Survey 1</th>
<th>Survey 2</th>
<th>Survey 3</th>
<th>Survey 4</th>
<th>Survey 5</th>
<th>Survey 6</th>
<th>Survey 7</th>
<th>Survey 8</th>
<th>Survey 9</th>
<th>Survey 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>38</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

Table 1: Appearance of water sample sequences in each survey, indicated by Y(Yes) and N(No). This Y/N was based off of any reads in any of the replicates (some had only very few reads). Survey 1 is blocked out because no water samples from this survey were sequenced.
<table>
<thead>
<tr>
<th>Survey</th>
<th>Haulout Site</th>
<th>Date Collected</th>
<th># Fecal sequences observed</th>
<th># eDNA sequences observed in same survey as matching fecal sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CH</td>
<td>6/23/2020</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>PT</td>
<td>6/25/2020</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>CH</td>
<td>7/6/2020</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>PB</td>
<td>7/6/2020</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>PH</td>
<td>7/21/2020</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>TI</td>
<td>8/3/2020</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>SI/SI-PLB</td>
<td>8/3/2020</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>SI</td>
<td>8/19/2020</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>9</td>
<td>LP</td>
<td>8/25/2020</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>TI</td>
<td>9/17/2020</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2: Comparison of the number of fecal samples and number of matching eDNA sequences observed in each survey. N/A under # Fecal sequences observed indicates that no fecal samples were collected in that survey, while N/A under # eDNA sequences observed in same survey as matching fecal sequences indicates that there were no matching eDNA sequences identified.
AUTHOR’S BIOGRAPHY

Sydney Jackson was born and raised in Southern California in her hometown of Upland, where she graduated from Upland High School in 2018. She began attending the University of Maine in the fall of that same year and plans to graduate in May 2022, with a major in marine science and a minor in journalism. During her time at UMaine, she was a member of the Alpha Omicron Pi sorority, a member and officer of the Tennis Club, a UMaine UVote Ambassador, and an opinion contributor for the school newspaper, The Maine Campus. In the summer of 2021, Sydney was hired as a student research assistant for the Maine-eDNA program, where she became more familiar with the collecting and filtering processes of the eDNA method. She then spent her senior year as a student research assistant in the Cammen Lab, working under a grant she was awarded by the Center for Undergraduate Research.

Upon graduation, Sydney plans to seek jobs in the scientific journalism field while also pursuing her passion of working with marine mammals.