Assessing the Effects of Whaling on the Genetic diversity of North Atlantic Blue Whales (*Balaenoptera musculus*)

By

Erin Stacey

Supervised By

Dr. Timothy Frasier

A Thesis Submitted To Saint Mary's University, Halifax, Nova Scotia in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science, Honours in Biology

April 2022, Halifax, Nova Scotia

Approved: Dr. Timothy Frasier

Supervisor

Approved: Dr. David Chiasson

Supervisor

Date: April 30, 2022

Table of contents	**Page Number**

Abstract	**3**
Introduction	**4**
Materials and Methods	**11**
Results	**19**
Discussion	**26**
References	**28**

List of Tables and Figures

Figure 1.The number of haplotypes for the ancient and contemporary samples. **19**

Table 1. Various measurements of genetic diversity were performed on the ancient samples using RStudio, Fabox, and calculations by hand. **19**

Table 2. The concentration of the DNA samples using the nanodrop. **21**

Table 3. The concentration of the diluted DNA samples using the nanodrop. **22**

Table 4. Various measurements of genetic diversity on the contemporary samples. **23**

Figure 2. The haplotype diversity values of the ancient and contemporary samples. **24**

Figure 3. The nucleotide diversity values of the ancient and contemporary samples. **25**

Abstract

Industrial whaling in the early to mid 20th century severely depleted the population of blue whales in the North Atlantic. Due to this reduction in population it is likely that the current population of blue whales in the North Atlantic have low genetic diversity. Blue whales are an essential part of the oceans ecosystem and are currently considered endangered. Our research looks to assess the levels of genetic diversity amongst the contemporary North Atlantic blue whale population compared with the ancient pre-whaling population. The ancient samples were collected by archaeologists from ancient Viking sites found in Iceland. The contemporary samples were collected from various blue whale carcasses that had washed ashore across the eastern coast. The mitochondrial DNA of the samples were extracted, amplified and sequenced. The DNA of the ancient samples was already extracted and sequenced. Various measurements of genetic diversity were calculated and compared between the ancient and contemporary samples. Preliminary findings suggest a decrease in genetic diversity of contemporary blue whales compared to their pre-whaling ancestors. Thirteen haplotypes were found to be unique to the ancient samples, one haplotype was found to be unique in the contemporary samples, and three haplotypes were shared between the ancient and contemporary samples. A haplotype is the group of alleles that is inherited together from one organism to another from a single parent. This makes haplotypes a useful tool for examining evolution and genetics of a species.

Introduction

Blue Whale Description

The blue whale, *Balaenoptera musculus,* is part of the order Cetartiodactyla (Hermosilla, 2015). Blue whales belong to the suborder Mysticeti, where baleen plates are used as a filtration feeding system. This differs from the other suborder, Odontoceti, where the marine mammals have teeth. Lastly, they are a part of the family Balaenopteridae, which are characterized by their pleated throats which facilitates their filter feeding. The pleats allow the whale to expand their throats to catch a large amount of water and euphausiids, where they then retract their throats pushing the water out through the baleen while leaving the euphausiids in the mouth to swallow (Clapham, 1999;Fisheries and Oceans Canada, 2009).

Blue whales are the largest known living animal on earth, with adults in the Antarctic being observed at lengths up to 33 m and weighing more than 150,000 kg (Clapham, 1999;Reeves et al., 1998). The Northern Atlantic blue whales are reported to be smaller though, with a maximum length of 27 m (Clapham, 1999;Reeves et al., 1998). They are slender and long in appearance compared to other Balaenopteridae whales. They have a broad and flat rostrum, where the rostrum is the front half of the cranium, consisting of the jaws, palate and nasal cavity (Rafferty et al., 2010). Blue whales mainly subsist on euphausiids (krill) and to feed they fill up their mouths with water and a school of krill which in turn expands their pleated throat and chest to create space for the water. The whale then pushes the water out of its mouth through the baleen which catches the krill and keeps it in the mouth of the whale. Once the water has been expelled from the body and the zooplankton have been swallowed, the body returns back to its original slender shape. Other unique characteristics of the blue whale are their smaller dorsal fin.

Compared to other Balaenopteridae the blue whale's fin is much smaller and placed further back, closer to the tail flukes rather than the midsection of the body (Clapham, 1999;Reeves et al., 1998). Their bodies are a mottled gray colour, with both light and dark gray splotches found throughout. The mottling on their backs are unique to each whale and serves as an identification system for scientists. Similar to a fingerprint, mottling patterns are discernable from other whales through colour, shape, and amount of mottling. Some whales have mottling all along their back with a mixture of gray shades while others may have hardly any mottling and could be limited in shade range (Clapham, 1999;Reeves et al., 1998).

Blue Whale Global Distribution

Blue whales are considered a cosmopolitan species, meaning their geographical distribution spans to most, if not all, of the regions around the globe. They are typically observed in a variety of habitats, from coastal waters to pelagic water far offshore (McDonald et al., 2006;Reeves et al., 1998). Although not much is known about their migratory behaviour it is believed that the blue whale partakes in seasonal migrations from high-latitude feeding grounds to breeding areas found in the tropics, but there is evidence of some blue whales occupying tropical waters all year round (McDonald et al., 2006;Reeves et al., 1998). The blue whales reside in feeding grounds during the summer for four to six months a year to feed in high krill populated areas to increase their body fat and prepare for the winter months where food is sparse compared to the feeding areas (McDonald et al., 2006;Reeves et al., 1998).

The blue whale species is divided into smaller groupings of subspecies. The number of subspecies of the blue whale is widely disputed as they are only vaguely differentiated by size

and geographical distribution. Currently, there are four accepted subspecies of blue whales in the scientific community. *B.m. intermedia* is the subspecies found in the Southern Ocean and Antarctic, *B.m. brevicauda,* also known as the pygmy blue whale, is found in sub-Antarctic in the Indo-Pacific Ocean as well as the south-eastern Atlantic Ocean, and *B.m. indica* is found in the northern Indian Ocean (McDonald et al., 2006;Reeves et al., 1998). The fourth and final subspecies, *B.m. musculus* is found in the Northern Hemisphere in both the Atlantic and Pacific Oceans (McDonald et al., 2006;Reeves et al., 1998). The North Atlantic population of blue whales will be the focus of this study.

The North Atlantic population can be further broken down into two smaller populations: the Northeast Atlantic population that inhabits waters around Iceland and the Northwest Atlantic that is found around eastern Canada (Fisheries and Oceans Canada, 2009).

Blue Whale Population Status

The Northwest Atlantic population of blue whales were designated endangered by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) in 2002 (Fisheries and Oceans Canada, 2009). They were also added to the *Species at Risk Act* (SARA) as an endangered species in 2005 (Fisheries and Oceans Canada, 2009). The current population size is unknown, but researchers estimate that there are less than 1000 individuals in the Northwest Atlantic, with some estimates of there being only around 250 mature blue whales (Fisheries and Oceans Canada, 2009).

Exploitation of Blue Whales Through Whaling

Exploitation of whales began hundreds of years ago through the act of whaling. Whaling involves hunting whales to use for food and oil and was performed by nations around the world. Due to the blue whales' speed and tendency to reside in deeper waters they were not big targets for the whaling industry until the beginning of the 20th century (Clapham, 1999). The invention of the steam engine allowed commercial whalers to attain speeds fast enough to keep up with the blue whales, and another new invention, the explosive harpoon, was an important tool used to hunt them. These two inventions were the main reason that blue whales were able to be caught, but other inventions such as the compressor, which was used to inflate the whale carcasses so they would not sink, and the factory ship were also major contributors to the exploitation of blue whales (Clapham, 1999). Furthermore, in 1904 a rich whaling ground business was opened in the Southern Ocean that contributed to the demise of blue whales. Blue whales were considered a desired whale by whalers due to its large size. As the largest whale and animal in the world they provide the most food and oil when caught. The International Whaling Committee worked to officially protect the blue whale and many other whale species worldwide in 1967 (Clapham, 1999). From the beginning of the century to the time the species was protected, approximately 360 000 blue whales were hunted and killed in the Antarctic alone (Clapham, 1999). In the North Atlantic 11,000 blue whales were estimated to have been captured prior to the ban in 1967 (Fisheries and Oceans Canada, 2009). From the years of 1898 to 1915, 1,500 blue whales were caught in eastern Canadian waters (Fisheries and Oceans Canada, 2009). Researchers believe the reduction in blue whales from pre-whaling to to now has been a staggering 70% decrease in population (Fisheries and Oceans Canada, 2009). Although whaling has been banned since 1967

by the IWC there has been no sign of recovery of the species in the North Atlantic (Clapham et al., 1999;Fisheries and Oceans Canada, 2009;Reeves et al., 1998).

Genetic Diversity and Bottleneck Effect

Genetic diversity within a population is essential for the conservation and management of endangered species. Populations with low genetic diversity become susceptible to extinction from threats that would otherwise not affect the status of a species (Bryant et al., 1986;Parra et al., 2018). Having low genetic diversity in a population limits their ability to withstand threats like disease and climate change as there is less differentiation among individuals, meaning there is a decreased number of haplotypes which can lead to the accumulation of negative genetic variations. With higher levels of genetic diversity populations have a higher chance of surviving threats since the threat may not affect all of the individuals in the population. For example, disease can be extremely deadly to a population with low genetic diversity since the population may all get sick from the disease whereas other populations with higher genetic diversity would have less chance of all the individuals being affected as they have more genetic diversity amongst them (Bryant et al., 1986;Parra et al., 2018).

The bottleneck effect is when a population undergoes a dramatic reduction of individuals and subsequent increase in population where the distribution of genetic variation among the population is severely affected (Bryant et al., 1986). When a bottleneck event occurs the individuals killed are completely random, meaning the surviving genes are also completely random. Certain alleles can be entirely lost which results in the loss of genetic diversity of the entire population as that allele can no longer be passed onto future generations. Furthermore,

bottlenecks have been proven to increase and amplify linkage, where linkage is a close association of genes and the closer they are to each other on the chromosome results in a higher chance that they will both be inherited. . Even if the population recovers in terms of number of individuals, the genetic diversity of that population will remain low for a long period of time – maintaining a signature of the bottleneck event (Bryant et al., 1986;Parra et al., 2018).

Mitochondrial DNA Importance in the Study

Mitochondrial DNA has become one of the key DNA markers of choice for the analysis of ancient DNA, because mitochondrial DNA molecules are often present even in situations where the nuclear DNA has become largely degraded – often to the point of not being usable for genetic analyses. Mitochondrial DNA is haploid, meaning there is only a single set of chromosomes. Furthermore, mitochondrial DNA is inherited from the mother to their offspring and no mitochondrial DNA is inherited from the father. This makes it easier for mitochondrial lineages to be traced than nuclear lineages which includes DNA from both the mother and father that has undergone recombination. Mitochondrial DNA also has a distinctive region which is different amongst individuals and provides a way to distinguish between individuals. This region is called the control region and will be the region amplified and focused on in this study. Mitochondrial DNA also has a simple genetic structure. This is important since complicated structures with features like repetitive DNA, transposable elements (sequences of DNA that may move location in the genome), and pseudogenes (regions of defective gene copies). Lastly, mitochondrial DNA evolves quickly, specifically the control region (Avise et al., 1987;Baker et al., 1990).

Additionally, mitochondrial DNA has benefits that are specific to this study. Mitochondrial DNA is much more abundant than nuclear DNA. Each mitochondria in a cell in the individual has its own supply of mitochondrial DNA, resulting in more copies of the mitochondrial DNA than nuclear DNA (Avise et al., 1987;Baker et al., 1990). This is beneficial for this study as contemporary blue whale DNA will be compared with ancient blue whale DNA found from old Viking sites. Having more copies means there is a higher chance that DNA will be successfully extracted from the ancient samples, whereas nuclear DNA would likely be too degraded to extract and amplify. Additionally, the shape of the mitochondrial genome makes it less prone to degradation than nuclear DNA. Specifically, the mitochondrial genome is circular: many of the enzymes that break down DNA require a "loose end" to start from (reference). However, the circular shape of the mitochondrial genome means that it is harder (and will take longer) for these enzymes to break down the mitochondrial DNA molecule (Avise et al., 1987;Baker et al., 1990).

Research Goal

Due to their low population numbers blue whales in the North Atlantic are at risk for extinction. Blue whales are important for the health and integrity of the ecosystems they reside in and effort should be made into the recovery and sustainability of the species. In the situation of the North Atlantic blue whale, industrial whaling may have been a bottleneck event that negatively impacted the genetic diversity of the species. This study aims to research the effects that whaling had on the North Atlantic population of blue whales by comparing the mitochondrial DNA of ancient blue whale samples collected from various archaeological Viking sites with mitochondrial DNA of contemporary blue whale samples collected from various carcasses that

had washed ashore. The goal of this thesis was to quantify the impact of whaling on the genetic diversity of North Atlantic blue whales. My approach was to compare sequence variation within the mitochondrial control region from contemporary samples and compare this to the same data from ancient blue whale samples collected from archaeological excavation of Viking sites. With further information on the genetics of blue whales and the impacts that industrial whaling had on the genetics of blue whales today will help in the process of recovery and conservation of the species.

Materials and Methods

Preparing blue whale tissue for extraction

Various types of blue whale tissue were subsampled to be extracted including skin and gum tissue. Skin is the typical tissue used in DNA extractions but gum tissue was also subsampled for extraction to compare the DNA yield to that of the skin samples. For the extraction approximately 40 mg of tissue was weighed and recorded. The tissue was then cut into small pieces with a razor blade to facilitate the breakdown process. The small pieces of tissue were then added to a 1.5 mL tube containing 400 μL of RCMP lysis buffer. The RCMP lysis buffer is a combination of 10 mM Tris (pH 8), 10 mM EDTA (pH 8), 2% SDS, 0.1 M NaCl, and 40 mM DTT, that had previously been pipetted into the tubes. The tube was then labeled with the specimen and repeated for each sample.

Next, the baleen was subsampled to be extracted. Baleen is more tough than skin or gum tissue samples so simply cutting it into small pieces would not work. The baleen was instead grated using a fine grater until approximately 50 mg of baleen was weighed in a weigh boat. The baleen

pieces were then added to a 1.5 mL tube containing 1000 μL of 0.5 M EDTA (pH 8). The samples were then left overnight to soak in the EDTA.

The next day the baleen samples were centrifuged at 14,000 x g for 2 minutes and the EDTA poured out while the baleen remained in the bottom of the tube. After the EDTA was poured out 400 μL of the RCMP lysis buffer was added to each sample and shaken until mixed thoroughly. At this time a negative control sample was made that has no sample in it but will be treated like every other sample going forward so that contamination can be found. The samples then sat at room temperature for 7 days mixing frequently each day, and then were frozen for 6 more days.

The next step in the breakdown process of the tissue is to add proteinase K. One of the main challenges in isolating DNA is removing all of the proteins that are attached to the DNA. Proteinase K breaks the protein down into smaller pieces which can then be removed from the DNA. 33.3 μL of proteinase K was added to each sample and mixed thoroughly. The samples then sat at room temperature overnight. The next day a second shot of 33.3 μL of proteinase K was added to each sample. The samples were then put in a water bath of 55 ℃ for one hour. Once the samples were removed from the water bath another shot of 33.3 μL of proteinase K was added to each sample. The samples were then incubated overnight at room temperature.

DNA extraction

Phenol:chloroform is the best extraction to use to get high quantities of DNA and is great for situations where there may not be a lot of DNA in the sample. The first step of the extraction is to add an equal volume (400 μL) of phenol:chloroform to each sample and upended by hand for 5 minutes. Next the samples were spun at 12,000 x g for 1 minute. While the samples were being spun 400 μL of phenol:chloroform was added to new labeled 1.5 mL tubes. Once the samples were done spinning they formed two layers, the bottom layer containing the cell debris and phenol:chloroform while the top aqueous layer contains the DNA and water. The top aqueous layer containing the DNA was then pipetted and transferred into the previously prepared new labeled tube with more phenol:chloroform. The new tubes were then upended by hand for 5 minutes and then spun at 12,000 x g for 1 minute. Next 400 μL of chloroform alone was added to new labeled 1.5 mL tubes. Once the samples were done spinning the top aqueous layer was then pipetted and transferred into the previously prepared new labeled tubes with just chloroform added to them. The samples were then upended by hand for 5 minutes and spun at 12,000 x g for 1 minute.

The next step of the DNA extraction process is to precipitate the DNA out of solution so that it can be dissolved in an appropriate volume. First, 80 μL of 10 M ammonium acetate was added to new labeled 1.5 mL tubes. The top aqueous layer of the spun samples was then added into the previously prepared tubes containing the 10 M ammonium acetate. Then, 800 μL of ice cold 95% ethanol was added to the samples, mixed well, and incubated overnight in the freezer at -20 ℃.

The last step of the phenol:chloroform extraction is to remove the salts and store the DNA in $TE_{0.1}$. The samples were fist spun at 12,000 x g for 10 minutes to ensure the DNA precipitate is packed tightly into a small pellet. Once the samples are done spinning the ethanol was decanted from the sample. This involves pouring out the ethanol from the tube but the pellet of DNA remains inside the tube. Next, 100 µL of 70% ethanol was added to each tube while rotating to rinse the sides of the tube to remove all salts. The samples were then spun again at 12,000 x g for 10 minutes. The 70% ethanol was then decanted and excess droplets were removed with a kimwipe. The tubes were then left open to air dry for 15 minutes to allow all of the ethanol to evaporate. Lastly, 100 µL of $TE_{0.1}$ was added to each sample and then the samples were put in a 55 ℃ water bath for 3 minutes to redissolve the pellet of DNA and mixed thoroughly.

DNA quantification

DNA quantification was estimated via spectrophotometry using a NanoDrop 1000 (ThermoFisher Scientific). Before quantification, each sample was mixed thoroughly to ensure even distribution of DNA in solution. The first sample to be done is the negative control as it contains no DNA and will create a blank to then be compared against for samples that do contain DNA. 2 µL of the sample was pipetted onto the nanodrop and the quantity was recorded.

PCR Amplification

To prepare the samples for PCR 50 µL aliquots of 5 ng/µL dilutions were created from the samples. Once the samples were diluted two sets of aliquots were created, one set for an agarose gel plate to assess the DNA and another for amplification PCR. Into two new labeled tubes 2 µL of sample was pipetted into each tube.

Before running the amplification PCR the cocktail was created with all the correct reagents. There were 22 samples, 2 positive control samples, 1 negative control sample, and 2 extra samples which came to a total of 27 samples with a desired volume of 20 µL per tube.

A PCR cocktail was then created to amplify the control region of each sample, using the primers t-Pro and Primer-2 from Yoshida et al. (2001). Specifically, the reaction conditions were as follows: 10ng of template DNA, 0.2nM of each dNTP, 1X PCR Buffer (Promega), 1.5 mM $MgCl_2$, 0.3µg/µl of BSA (bovine serum albumin), 0.3mM of each primer, and 0.05U/µl of *Taq* DNA polymerase (Promega). The cycling conditions were then as follows: An initial denaturation time of 5 minutes at 94ºC; 30 cycles of 94ºC for 30 second, 55ºC for 1 minute, and 72ºC for 1 minute; and a final extension step of 60ºC for 45 minutes.

Visualizing PCR Products

To visualize how well each sample amplified, and therefore to determine which samples to proceed with for sequencing, and what the relative concentration of PCR products were for each, PCR products were size-separated by electrophoresis through a 1.5% agarose gel and visualized with ethidium bromide. 2 µL of Orange-G dye was pipetted into new tubes then 5 µL of sample was pipetted into the new tubes with orange-G dye. To make the gel 80 mL of 0.5x TBE was poured into a flask. Then 1.2 g of agar powder was added to the flask and microwaved for 3 minutes, carefully stirring every 30 seconds. Once removed from the microwave 4 µL of ethidium bromide was added to the flask. The mixture was then allowed to cool slightly and then poured into the mold and allowed to set. To run the gel 7 µL of the orange-G dye aliquots were

pipetted into individual wells on the gel plate. Into the first well 5 μL of the size standard was pipetted into both the top and bottom rows of the gel. The gel was then run to visualize the DNA and prepare for the sequencing reaction.

DNA Sequencing

The sequencing reactions in our laboratory yield the best results when approximately 10 ng of DNA is used for every 100 base pairs (bp) of desired sequence. The primers that were used amplified a ~400 bp fragment of the control region, and therefore ~40ng of PCR product was used for each sample in the subsequent sequencing reaction.

However, prior to the sequencing reaction, unused primers and dNTPs must first be removed. I did this using an enzymatic approach. Specifically, based on the brightness of the PCR products from the agarose gel (relative to the Low-Mass DNA Ladder that was also loaded onto the gel), the concentration of PCR products was estimated for each sample. The appropriate volume was then aliquoted into new tubes, so that there was roughly 40 ng of product for each sample. A cocktail was then made of the appropriate enzymes to break down the unused primers and dNTPs in each sample. Specifically, for each sample this consisted of adding 0.65 μL of Antarctic Phosphatase Buffer, 0.1 μL of Antarctic Phosphatase, and 0.03 μL of ExoNuclease I (enzymes and buffers from New England Biolabs). The samples were then incubated at 37℃ for 15 minutes, and the enzymes then denatured by incubation at 80℃ for 15 minutes.

The sequencing reaction was then conducted by making a cocktail of the following reagents, and then adding 9.22 μL to each samples: 0.25X Sequencing Reaction Mix (Applied Biosystems),

1X Sequencing Buffer (Applied Biosystems), and 0.33 mM of the Primer-2 primer. The cycling conditions were then as follows: An initial denaturation period at 96℃ for 2 minutes; 30 cycles of 96℃ for 20 seconds, 50℃ for 20 seconds, and 60℃ for 4 minutes.

The next steps were performed by Timothy Frasier and included ethanol precipitation to de-salt the samples, adding 2 µL of DNA to 10 µL of formamide for the visualization process, and then separating the samples by size and visualizing them on an ABI 3500XL Genetic Analyzer.

Analysis of Sequences

The program MegaX (MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms (Kumar, Stecher, Li, Knyaz, and Tamura 2018)) was used to visualize the sequence files, and each sequence was edited by trimming away messy portions at the beginning and end of each sequence, as well as by scanning through each peak of each electropherogram to ensure that the bases were called correctly. After editing, the sequences were aligned in MegaX using the method implemented with Clustal (Thompson et. al., 1994).

The online tool FaBox (Villesen, 2007) was then used to identify the variable sites across the sequences, as well as to identify which samples had which haplotype. This data provided the following metrics of diversity: (1) number of variable sites; (2) number of haplotypes; and (3) nucleotide polymorphism, which is calculated as the number of variable sites divided by the number of sites examined.

The other metrics of genetic diversity of importance were haplotype diversity and nucleotide diversity. Haplotype diversity is the probability that two randomly sampled individuals will have

a different haplotype. Thus, populations dominated by one or a few common haplotypes will have very low haplotype diversity, whereas populations with a large number of haplotypes will have higher haplotype diversity. Nucleotide diversity is the average *difference* between the haplotypes of two randomly selected individuals (including the probability that they have the same haplotype). Thus, nucleotide diversity considers both the probability of two haplotypes being different, as well as the magnitude of that difference. Estimates for these metrics were obtained using the hap.div and nuc.div functions of the *pegas* R package, respectively (Paradis, 2010).

Determining Number of Individuals of the Ancient Samples

For the ancient samples it was impossible to determine which bone fragments came from which individual. This means that two bone fragments could be from the same individual. This could skew the results if multiple individuals were counted more than once since it would give inaccurate calculations to the actual population. Instead, two measurements were taken with the minimum number of possible individuals and the maximum number of possible individuals. This was done by examining an excel sheet containing information like the haplotype and location where the bones were found. For the minimum number of possible individuals only one sample was counted if it had the same haplotype and was found in the same location. For the maximum number of possible individuals every sample was counted as its own individual.

Results

For the ancient samples there were two data sets, a "minimum" and "maximum" number of individuals. The minimum number of individuals consisted of 33 individuals, while the maximum number of individuals consisted of 55 individuals (**Table 1**). For the ancient samples, a total of 16 haplotypes were identified in both the minimum and maximum number of individual sample sets (**Figure 1**).

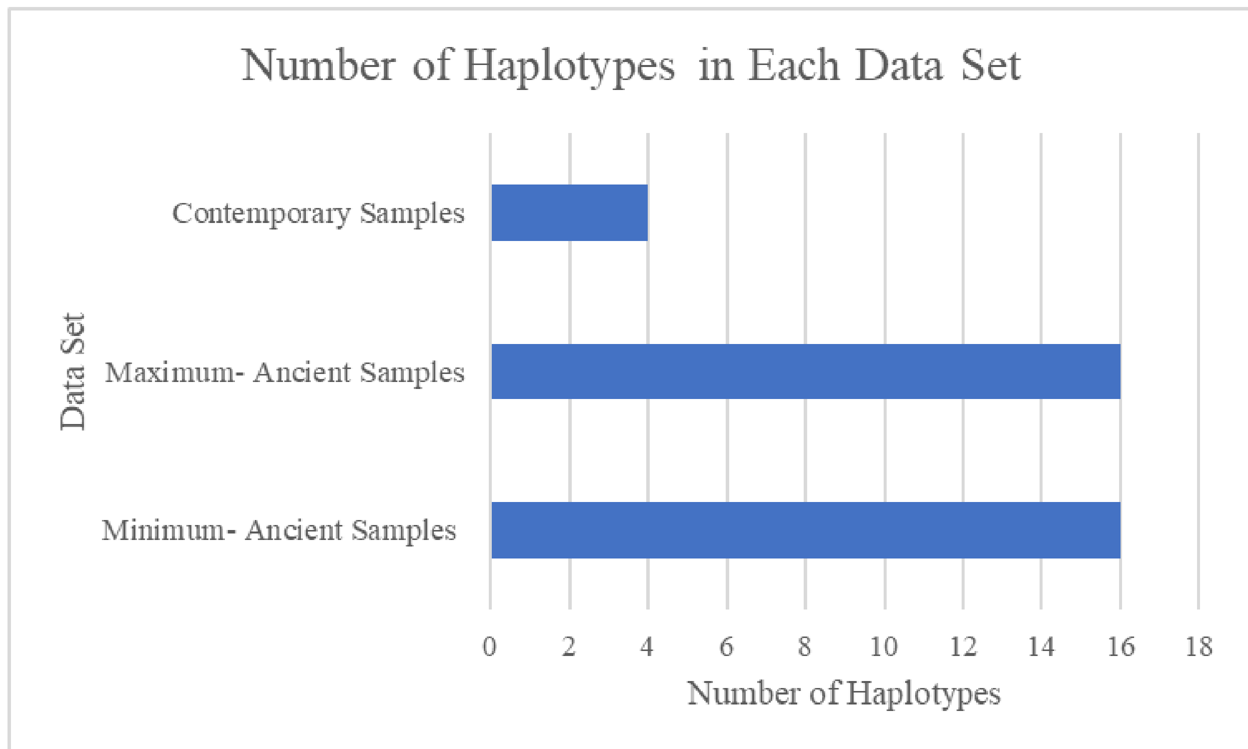

Figure 1. The number of haplotypes for the ancient and contemporary samples.

Table 1. Various measurements of genetic diversity were performed on the ancient samples using RStudio, Fabox, and calculations by hand.

	Minimum Number of Individuals- Ancient Samples	Maximum Number of Individuals- Ancient Samples

Number of Variable Sites	17	17
Number of Haplotypes	16	16
Sequence Length	421	421
Nucleotide Polymorphism	0.040	0.040
Haplotype Diversity	0.93	0.91
Nucleotide Diversity	0.0070	0.0071
Number of Individuals	33	55

Before the samples could be sequenced they needed to be quantified using the nanodrop (**Table 2, Table 3**). This gave the concentration of the DNA which was used to calculate required amounts of DNA in later steps.

For the contemporary samples 20 samples were sequenced which consisted of 11 individuals, a much smaller sample size than the ancient samples (**Table 4**). Of the 11 individuals, 10 were blue whales but one was not. It was discovered that this whale was a hybrid between a fin whale and a blue whale. Therefore, there were 10 blue whale individuals. For the contemporary samples, a total of 4 haplotypes were identified using only the one contemporary sample set as we know the number of individuals (**Figure 1**).

Table 2. The concentration of the DNA samples using the nanodrop.

Sample ID	**ng/µL**
BMU 0001	124.9
BMU 0007	22.7
BMU 0010	49.0
BMU 0011	169.0
BMU 0012	289.0
BMU 0013	91.6
BMU 0014	108.7
BMU 0015	105.1
BMU 0018	81.7
BMU 0027	163.6
BMU 0028	57.6

Some samples had very high amounts of ng/µL which would give inaccurate readings. To fix this any sample above 250 ng/µL was diluted in a 10:1 ratio with the $TE_{0.1}$ solution and redone on the nanodrop (**Table 3**).

Table 3. The concentration of the diluted DNA samples using the nanodrop.

Sample ID	**ng/µL**
BMU 0002	51.2
BMU 0004	28.4
BMU 0005	30.7
BMU 0006	107.8
BMU 0008	162.6
BMU 0009	44.8
BMU 0023	123.0
BMU 0025	50.2
BMU 0026	87.0

Table 4. Various measurements of genetic diversity were performed on the contemporary samples using RStudio, Fabox, and by hand.

	Contemporary Samples
Number of Variable Sites	10
Number of Haplotypes	4
Sequence Length	402
Nucleotide Polymorphism	0.025
Haplotype Diversity	0.71
Nucleotide Diversity	0.0075
Number of Individuals	10

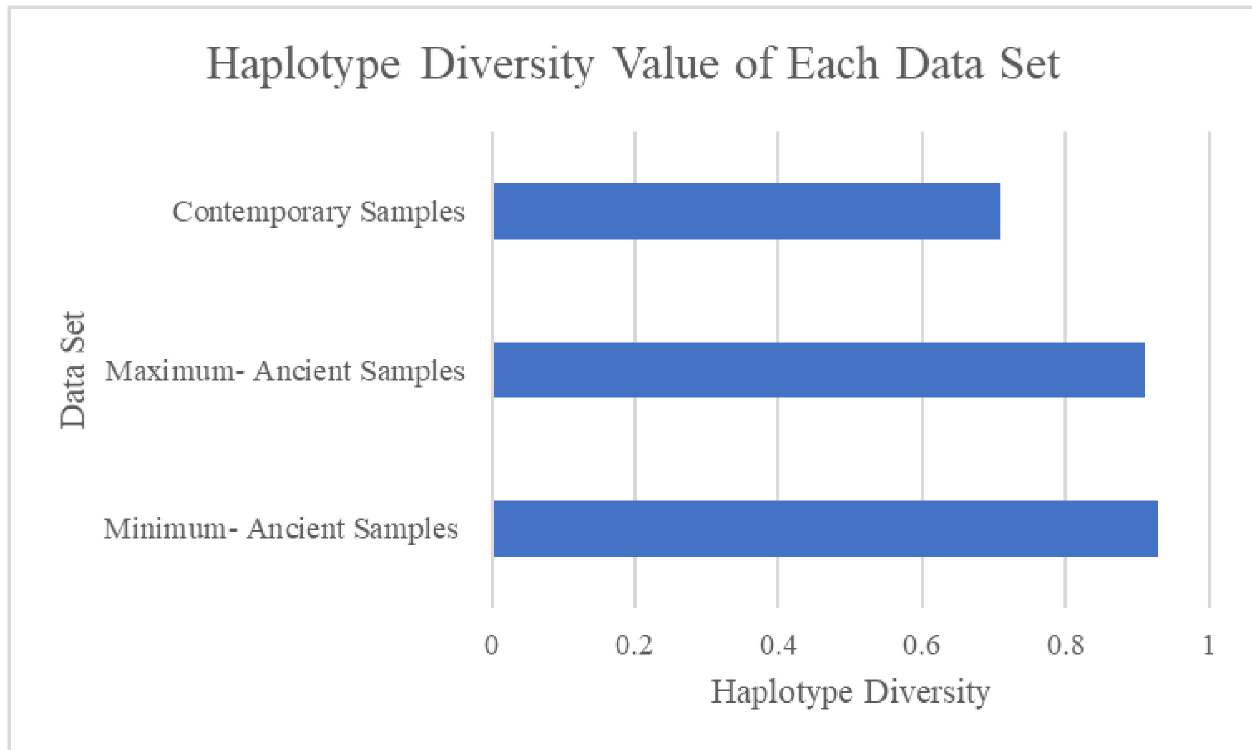

Figure 2. The haplotype diversity values of the ancient and contemporary samples.

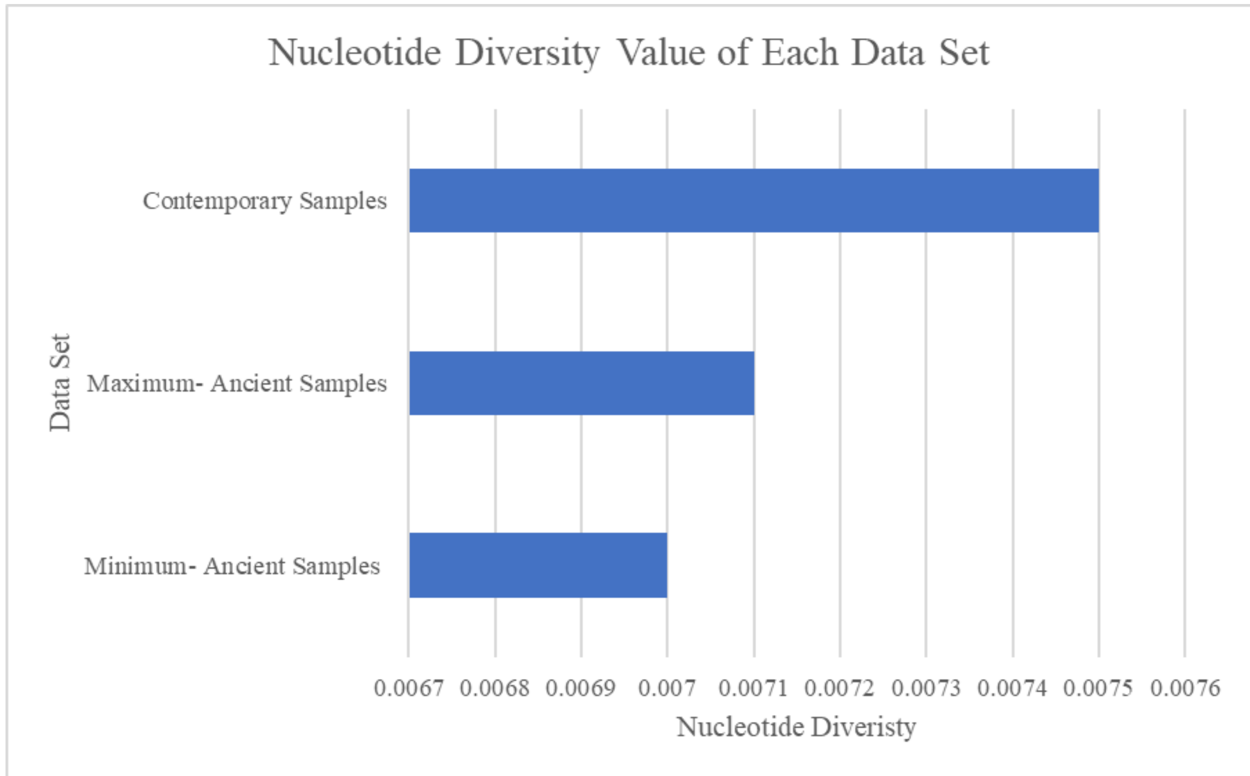

Figure 3. The nucleotide diversity values of the ancient and contemporary samples.

For the ancient samples the haplotype diversity in the minimum number of individuals was 0.93 and for the maximum number of individuals was 0.91. As for the contemporary samples, the haplotype diversity was 0.71, a decrease from the ancient samples (**Figure 2**).

For the ancient samples the nucleotide diversity in the minimum number of individuals was 0.0070 and for the maximum number of individuals was 0.0071. As for the contemporary samples, the nucleotide diversity was 0.0075, a slight increase from the ancient samples (**Figure 3**).

Discussion

Knowledge about blue whales, primarily their population structure, is severely lacking due to their elusive nature. Amongst the scientific community it is widely accepted that the exploitation of blue whales led to diminished populations, but quantitative data on the populations of blue whales is scarce (Clapham, 1999). With new technology advancements, it is becoming more accessible to study blue whales. Studies such as this one is the beginning to learning the unknown about blue whales and their populations.

When beginning to examine the sequences it was clear that one of the sequences was very different from the rest. This raised some questions, the sequence was then compared on GenBank, a collection of genetic sequences available to the public. The results of this search showed that the sample did not belong to a blue whale but instead had the mitochondrial sequence of a fin whale. Each whale when found had a necropsy performed on them, meaning this whale was identified as a blue whale by the field team. This means that this whale was likely a hybrid between a blue whale and a fin whale. In this case the hybrid whale would have a fin whale mother and a blue whale father since the mitochondrial DNA was that of a fin whale. It is theorized that these hybrids came to be due to reduced population sizes in closely related cetaceans (Pampoulie et. al., 2020). Fin whales and blue whales are the two largest animals on earth. The main difference between a fin whale and a blue whale is that the fin whale is smaller than the blue whale. Furthermore, fin whales are often sighted alongside blue whales. Fin whales and blue whales are commonly seen in mixed schools together, one paper reported from 1981 to 1987 having nine sightings of blue whales and five of those sightings were accompanied with fin whales (Berube & Aguilar, 1998). Oftentimes in cases of hybrids between species there are many

conditions associated with it. For example, the species of the mother and father are often strict. In the case of the fin whale and blue whale though it appears to not matter which species is the mother and which species is the father. It is possible to have a fin whale mother with a blue whale father like in this case, but it is also possible to have a fin whale father and a blue whale mother (Arnason et. al., 1991). Although this hybridization is known to occur between fin whales and blue whales and is therefore not very surprising, it was still an interesting finding.

In total the ancient samples had 16 different haplotypes while the contemporary samples had 4 different haplotypes. Of these haplotypes 13 of them were unique to the ancient samples, 1 haplotype was unique to the contemporary samples, and 3 haplotypes were shared between the ancient samples and contemporary samples. This is a dramatic decrease in the different haplotypes from ancient to contemporary, indicating that the levels of genetic diversity have decreased. Furthermore, only one of the haplotypes was unique to the contemporary samples so it is possible that the other haplotypes were lost through whaling. It is important to keep in mind the sample sizes of the ancient samples and contemporary samples while comparing them. For the contemporary samples there were only 10 individuals while the ancient samples had between 33 to 55 individuals. This is a large disparity between the two and would majorly affect the comparison. The 16 haplotypes of the ancient samples were across 33 to 55 individuals while the 4 haplotypes of the contemporary samples were across 10 individuals. Even if genetic diversity was exactly the same it is likely because of the lower number of individuals there will be less haplotypes. Further research would have to be performed with larger sample sizes of the contemporary samples to accurately estimate if genetic diversity has changed due to whaling.

References

Arnason, D., Spilliaert, R., Palsoottir, A. & Arnason, A. 1991. Molecular identification of hybrids between the two largest whale species, the blue whale (*Balaenoptera musculus*) and the fin whale (*B. physalus*). - Hereditas JJ5: 183-189. Lund, Sweden.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. (1987). Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and Systematics. *Annual Review of Ecology and Systematics*, *18*(1), 489–522. https://doi.org/10.1146/annurev.es.18.110187.002421

Baker, C., Palumbi, S., Lambertsen, R., Weinrich, M., Calambokidis, J., O'Brien, S. (1990). Influence of seasonal migration on geographic distribution of mitochondrial DNA haplotypes in humpback whales. *Nature* 344, 238–240. https://doi.org/10.1038/344238a0

Berube, M., & Aguilar, A. (1998). A new hybrid between a blue whale, *Balaenoptera musculus*, and a fin whale, *B. Physalus*: Frequency and implications of hybridization. *Marine Mammal Science*, *14*(1), 82–98. https://doi.org/10.1111/j.1748-7692.1998.tb00692.x

Bryant, E. H., McCommas, S. A., & Combs, L. M. (1986). The effect of an experimental bottleneck upon quantitative genetic variation in the Housefly. *Genetics*, *114*(4), 1191–1211. https://doi.org/10.1093/genetics/114.4.1191

Clapham, P. J., Young, S. B., & Brownell, R. L. (1999). Baleen whales: Conservation issues and the status of the most endangered populations. *Mammal Review*, *29*(1), 37–62. https://doi.org/10.1046/j.1365-2907.1999.00035.x

Erich Hoyt. (2016). Blue Whale (Balaenoptera musculus). *The Eastern Arctic Seas Encyclopedia*, 54–55. https://doi.org/10.1007/978-3-319-24237-8_82

Fisheries and Oceans Canada, 2009. Recovery Strategy for the blue whale (*Balaenoptera musculus*), Northwest Atlantic population, in Canada [PROPOSED]. Species at Risk Act Recovery Strategy Series. Fisheries and Oceans Canada, Ottawa. 62 pp.

Hermosilla, C., Silva, L. M. R., Prieto, R., Kleinertz, S., Taubert, A., & Silva, M. A. (2015). Endo- and ectoparasites of large whales (cetartiodactyla: Balaenopteridae, Physeteridae): Overcoming difficulties in obtaining appropriate samples by non- and minimally-invasive methods. *International Journal for Parasitology: Parasites and Wildlife*, *4*(3), 414–420. https://doi.org/10.1016/j.ijppaw.2015.11.002

McDonald, M. A., Hildebrand, J. A., & Mesnick, S. (2009). Worldwide decline in tonal frequencies of Blue Whale Songs. *Endangered Species Research*, *9*, 13–21. https://doi.org/10.3354/esr00217

Pampoulie, C., Gíslason, D., Ólafsdóttir, G., Chosson, V., Halldórsson, S. D., Mariani, S., Elvarsson, B. Þ., Rasmussen, M. H., Iversen, M. R., Daníelsdóttir, A. K., & Víkingsson, G. A. (2020). Evidence of unidirectional hybridization and second‐generation adult hybrid between the two largest animals on Earth, the fin and Blue Whales. *Evolutionary Applications*, *14*(2), 314–321. https://doi.org/10.1111/eva.13091

Paradis E (2010). "pegas: an R package for population genetics with an integrated–modular approach." *Bioinformatics*, **26**, 419-420.

Parra, G. J., Cagnazzi, D., Jedensjö, M., Ackermann, C., Frere, C., Seddon, J., Nikolic, N., &

Krützen, M. (2018). Low genetic diversity, limited gene flow and widespread genetic bottleneck effects in a threatened dolphin species, the Australian humpback dolphin. *Biological Conservation*, *220*, 192–200. https://doi.org/10.1016/j.biocon.2017.12.028

Rastogi, T., Brown, M. W., McLeod, B. A., Frasier, T. R., Grenier, R., Cumbaa, S. L., Nadarajah, J., & White, B. N. (2004). Genetic analysis of 16th-century whale bones prompts a revision of the impact of Basque whaling on right and bowhead whales in the western North Atlantic. *Canadian Journal of Zoology*, *82*(10), 1647–1654. https://doi.org/10.1139/z04-146

Reeves, R. R., Smith, T. D., Josephson, E. A., Clapham, P. J., & Woolmer, G. (2004). Historical observations of Humpback and blue whales in the North Atlantic Ocean: Clues to migratory routes and possibly additional feeding grounds. *Marine Mammal Science*, *20*(4), 774–786. https://doi.org/10.1111/j.1748-7692.2004.tb01192.x

Sudhir Kumar, Glen Stecher, Michael Li, Christina Knyaz, and Koichiro Tamura (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*(22), 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Torres‐Florez, J. P., Hucke‐Gaete, R., Rosenbaum, H., & Figueroa, C. C. (2014). High genetic diversity in a small population: The case of C Hilean Blue Whales. *Ecology and*

Evolution, *4*(8), 1398–1412. https://doi.org/10.1002/ece3.998

Villesen, P. (2007). FaBox: An online toolbox for fasta sequences. *Molecular Ecology Notes*, *7*(6), 965–968. https://doi.org/10.1111/j.1471-8286.2007.01821