Learning from the past: Genetic analyses of ancient and contemporary samples identify how historic and pre-historic events have shaped modern whale populations

By
Stephanie Louise Béland

A Thesis Submitted to
Saint Mary's University, Halifax, Nova Scotia
in Partial Fulfillment of the Requirements for
the Degree of Master of Science in Applied Science.

August, 2014, Halifax, Nova Scotia

Approved:	Dr. Timothy Frasier Supervisor
Approved:	Dr. Zhongmin Dong Supervisory Committee
Approved:	Dr. Tony Charles Supervisory Committee
Approved:	Dr. Marina Silva-Opps External Examiner
Date:	August 22, 2014

Abstract

Learning from the past: Genetic analyses of ancient and contemporary samples identify how historic and pre-historic events have shaped modern whale populations

by Stephanie Louise Béland

The techniques for genetic species identification and inferring past population sizes are being utilized in more fields than just population genetics. Anthropology, conservation biology, and species management are also using these practices. They can use these techniques to confirm historical data and to make informed decisions in the future. By using genetic species identification on ancient whale bones recovered from traditional First Nations whaling sites, it was discovered that gray and humpback whales were the species of choice. This is important information in regards to the future of First Nations traditional whaling on Canada's west coast. The DNA recovered from these ancient whale bones was then used to estimate pre-commercial whaling genetic variation, and make inferences about historic and pre-historic population demography. In addition to gaining insight into historic population demography, I also found that results differed dramatically if such inferences were based on historic or contemporary samples. These results are particularly important given the increasing popularity of using contemporary samples to infer population history, and show that caution is required when interpreting the results of such studies.

August 22, 2014.

Acknowledgements

I would first like to thank Timothy Frasier for the opportunity to expand my honours thesis and to continue working on a project that I am very passionate about. I would also like to thank him for all his guidance both in and out of the lab. Secondly, I would like to thank Jim Darling for all his help with sample collection, guidance, and collaboration over the last few years. A special thank you to the Tla-o-qui-aht First Nations and Joe Martin as this project would not have been possible without their support. I would also like to thank the members of the Frasier Lab, particularly Brenna, Krista and Lynne, for all their assistance and input over the course of this project. For funding I would like to thank Saint Mary's University, Faculty of Graduate Studies, and the Canadian Wildlife Federation.

A very special thank you to all those that I have met throughout my time here at St. Mary's University. My family, particularly my parents, for all their support when I told them I was moving across the country for this opportunity, and especially my roommate and good friend Laura Johnson, for all her support and keeping me sane. Last, but not least, my fellow graduate students for all their kinds words and encouragement.

Table of Contents

List of Figures:

estimated genealogies, and blue is the highest posterior density interval, which can be thought of as the confidence interval.

List of Tables

Chapter 2:

Chapter 3:

Chapter 1

General Introduction

Whaling on Canada's West Coast

Oral history, archaeological evidence, and some written accounts by early Europeans that made their way to North America, have suggested First Nations whaling traditions on the west coast of Canada date back several hundred to thousands of years ago (Huelsbeck 1988; Clayton 2000; Monks *et al.* 2001; Coté 2010). The Nuu-chah-nulth First Nations is a larger group that encompasses 14 smaller groups along Vancouver Island's west coast and they are an example of those that once practiced traditional sustenance whaling (Clayton 2000; Monks *et al.* 2001; McMillan & Claire 2005; Coté 2010). The Nuu-chah-nulth belong to the southern branch of the West Coast First Nations' Wakashan language family (Clayton 2000; Coté 2010). The Makah, who live on the Olympic Peninsula of Washington State in the USA, are also another member of this language family, and along with shared language, the Nuu-chah-nulth and Makah also share similar cultural patterns and the tradition of hunting whales (Monks *et al.* 2001; Coté 2010). Some anthropologists suggest that the Makah were once from Vancouver Island and moved to Washington State to resettle before European contact (Coté 2010).

The dramatic increase of commercial whaling, and building of processing plants along the west coast of North America, eventually lead to the decline in all whale species in the Pacific Ocean and put an end to sustenance whaling in the area (Coté 2010). Many First Nations groups were aware of the great reduction in whale population sizes and voluntarily stopped their whaling practices by the early 1900s (Reeves 2002; Coté 2010). The decline in whale numbers was also noted by the commercial whaling industry, but it was not until the International Whaling Commission (IWC), in cooperation with many countries around the globe, placed a moratorium on commercial whaling in 1986 did the

rapid decline in whales eventually stop (Coté 2010; Frasier *et al.* 2011). Since this moratorium, many whale species found in the Pacific Ocean have seen an increase in their population size, with some being delisted from endangered species lists in Canada and the United States (Coté 2010; Frasier *et al.* 2011). But this disruption in traditional whaling has lasted many decades for the First Nations of western North America, and with the increase in whale numbers the Nuu-chah-nulth and the Makah are now in similar situations when it comes to the issue of resuming whaling practices. The Makah are better known in the media in regards to their fight for their right to resume this cultural practice, for when they killed their first gray whale after at least a 75 year hiatus, it was not without great controversy and opposition from many conservation groups and the public (Cernetig 1997; Coté 2010; Verhovek 1999a; Verhovek 1999b). In addition, even though the Makah were given to resume their hunt, it has since been put on hiatus due to litigation associated with which governing bodies have the authority to grant such permission (Coté 2010).

The Nuu-chah-nulth have not attempted to take any whales in accordance to their cultural practices, but a subgrouping of five of the fourteen, consisting of the Huu-ay-ah, Kyuquot/Cheklesaht, Toquaht, Uchucklesaht and Yuu-cluth-aht, have recently negotiated a treaty that includes their right to take whales with the province of British Columbia, and Government of Canada, which came into effect in April of 2011 (Coté 2010). With this legal right to resume whaling, there is an increasing necessity to ensure that the populations they target have recovered to their pre-whaling estimates and to a level that can support hunting pressures (Coté 2010). Obtaining such information is complicated by the fact that there is some discrepancies between the oral history, archaeological

evidence, and literature regarding which species were pursued. Clarifying this issue requires two main investigations: 1) identification of the whale species that were traditionally targeted by the Nuu-chah-nulth First Nations and in what proportions and, 2) examining the demographic history of the targeted species to ensure that appropriate management goals are outlined to safeguard the survival of the population.

Species Identification and Ancient DNA

Evidence suggests that the whaling traditions on the West coast of Vancouver Island date back thousands of years, and therefore the DNA from any recovered whale bones is likely of low quality and quantity, and vulnerable to contamination from contemporary DNA of the same species. This is because once an organism dies, so begins the process of DNA degradation. Nucleases, which are enzymes inside cells that break down DNA, start to cleave the DNA into smaller and smaller fragments as soon as the cell dies. Additionally, during physical decomposition of the organism, the DNA is digested by micro-organisms, and damaged by the environment (Allentoft *et al.* 2012). The rate at which DNA breaks down is dependent on the environment, with hot and humid being the worst conditions for the preservation of DNA, and cold and dry being the best (Lindahl 1993; Mulligan 2006; Foote *et al.* 2012). Allentoft et al. (2012) estimated that a 500 base pair fragment of mitochondrial DNA, if kept at -5°C, would have a half-life of 9500 years. This colder temperature is ideal, but does not represent the conditions of many ancient specimens found in warmer substrates, in which DNA degradation will be faster.

This rate of degradation means that DNA recovered from old specimens (also called ancient DNA or aDNA) is often of both low quantity and quality, and therefore

special protocols are needed to for analyses to be successful. Protocols for working with aDNA include, but are not limited to: 1) minimizing cross-contamination between samples by wearing, and changing, the appropriate equipment such as gloves, and protective suits, if necessary, between handling of samples (Yang & Watt 2005); 2) ensuring that all bench space and equipment is adequately cleaned with bleach between samples to prevent cross-contamination (Yang & Watt 2005); and 3) preforming all pre-amplification work in an isolated laboratory in which no analyses of modern samples have been conducted (Cooper & Poinar 2000; Yang & Watt 2005; de Bruyn *et al.* 2011). Even when all the necessary protocols are followed, there is still a chance that no DNA will be recovered, and this may be due to the environment in which the sample was found or even how it was stored once removed from its original collection area. For example, it has been suggested that degradation could intensify when the samples are removed from their deposition environment, particularly if the laboratory environment is warmer and/or more humid than the original environment (Pruvost *et al.* 2007).

To overcome these obstacles, aDNA analyses are often performed using mitochondrial DNA (mtDNA), which is an important tool for species identification, but particularly when performing analyses on aDNA, because its structure and abundance increase the likelihood of obtaining large fragments of DNA (Linacre & Tobe 2011). The mitochondria organelles are found within the cells of eukaryotes and they are responsible for many functions, the most important being cellular respiration (Campbell & Reese 2008). Therefore, the number of mitochondria per cell is dependent on the cell's level of metabolic activity, therefore there can be anywhere from a single mitochondrion to thousands (Campbell & Reese 2008). There are multiple copies of mtDNA per cell, as

opposed to only the 2 copies of nuclear DNA (nDNA) per cell, because there are multiple mitochondrion within each cell and each have their own copy of DNA (Campbell & Reese 2008; Linacre & Tobe 2011). Mitochondrial DNA is a double-stranded circular molecule that is found within the mitochondria organelles (Budowle *et al.* 2003). The circular structure of mtDNA is helpful against breakdown as there are no exposed "ends" that are accessible to exonuclease activity (Allentoft *et al.* 2012). An initial prevention to mtDNA degradation is the mitochondrion's protein coat, which can protect the organelle from the destructive effects of enzymes (Karp 2005; Linacre & Tobe 2011). So this factor, in addition to its circular structure, makes mtDNA more resistant to degradation from extreme environmental conditions than nuclear DNA (nDNA) (Butler & Levin 1993).

The cytochrome *b* gene of the mtDNA is frequently used for species identification for taxonomic and forensic purposes (Irwin *et al*. 1991; Parson *et al*. 2000; Linacre & Tobe 2011). This gene codes for a protein involved in the electron transport chain of the mitochondria, so its function is critical to the survival of the organelle and the organism. Though mutations are possible within this gene, those changes that disrupt its overall functionality (i.e., that disrupt the coding of the necessary protein) will be selected against, as they may not allow the organism to survive or reproduce (Linacre & Tobe 2011). Because most mutations within this gene will cause a change in function, and therefore be selected against, little variation is found within populations. However, neutral mutations will slowly accumulate, resulting in differences between species. This characteristic of little within-species variation, but substantial among-species differences makes the cytochrome *b* gene useful for species identification (Tobe *et al*. 2010).

The control region, also called the displacement loop (D-loop), of the mtDNA is the site where replication of the mtDNA begins, but is largely a non-coding region so mutations are more frequent and seen more readily because there is less selection pressure against them (Larizza *et al*. 2002). This region, on its own, is not often used for species identification, at least between highly diverged species, because the high mutation rate can result a reduction in signal over time. For example, certain sites along the control region have very high mutation rates. However, only four bases can occur in a DNA sequence, and therefore if multiple mutations occur at the same site it can "erase" the previous mutation, making two sequences look more similar than they really are. This type of mutation, where a new mutation erases an old one, is called homoplasy, and they become more common as the two sequences being compared become more distantly related. However, when control region sequences are combined with those of the cytochrome *b* gene, species identifications are reliable (Yang & Speller 2006). This is due to the cytochrome *b* gene being able to distinguish between distantly related species, such as the different members of the Class Mammalia, and the control region for more closely related species such as those within the Suborder Mysticeti.

Moreover, the control region sequences are useful for estimating the minimal number of individuals taken. For example, bones found at a site could be from one, or many whales. The low variability of the cytochrome *b* gene means that many different individuals will have the same sequence, and thus it would not be possible to tell of multiple bones came from the same, or different, individuals. However, the higher variation of the control region means that most individuals will have different sequences, and therefore the control region is useful for estimating the minimum number of

individuals represented in a sample set, rather than just the number of bones.

Markers for Molecular Analyses

By using molecular species identification methods on whale bones from Vancouver Island, it is possible to discover what species were hunted, and what the relative proportions of different species were in the catch. This is important for cultural and biological reason. Knowing what species where historically taken can assist in accurately conveying the traditions of the Nuu-chah-nulth to future generations and the rest of Canada. We can also uncover if there were geographical differences in whale preference by comparing these findings to data from areas of Washington State, where the Makah traditionally hunted gray whales. This will be useful in discovering which species were included in traditional hunts, and thus should be the proposed targets if traditional whaling were to resume.

The sequence data obtained through species identification will also be useful for inferring the demographic history of these whale species before their populations were dramatically reduced by the commercial whaling industry (Rooney *et al.* 2001; Foote *et al.* 2012; Alter *et al.* 2012). For these analyses, the sequences of the control region will be more informative because their higher substitution rate will track historic demographic changes more readily than the slower-mutating cytochrome *b* gene.

A second set of molecular markers that will be utilized in investigating the demographic history of the whale species are microsatellites, also called short tandem repeats. Microsatellites are portions of the nuclear DNA that are made of 2-8 tandemly repeating nucleotides (Weber & Wong 1993; Pompanon *et al.* 2005). During

replication, microsatellites experience high levels of mutation due to "DNA slippage", resulting in increased variation on the number of tandem repeats (Bennett 2000). Microsatellites are useful for inferring recent historic demography due to their higher mutation rate, with an average rate from 10^{-4} to 10^{-3} per locus per generation (Weber & Wong 1993). This mutation rate allows researchers to get a recent picture of the past demography, whereas mtDNA, due to its slower mutation rate, is more useful for pre-historic demography (Veeramah & Hammer 2014). These characteristics of microsatellites and mtDNA are the reasons they are used together to get a complete picture.

Microsatellites will also provide more resolution for obtaining an estimation of the minimum number of individuals in the bone samples (Campbell & Reese 2008). These multilocus genotypes are widely used in population genetic studies to investigate population structure, changes in population size and gene-flow, because within a population a microsatellite locus can have multiple alleles and they can be found at varying frequencies (Pompanon *et al.* 2005; Veerameh & Hammer 2014). We can use these frequencies and their changes over time to make inferences about the population in question.

Inferring Historic and Pre-Historic Demography

One issue when trying to assess the status of a species for conservation or management purposes is estimating what the "natural" state of the population was (i.e., the status of the population prior to the impact of humans). This information is often very difficult to obtain, but is necessary for identifying the recovery status of a population,

setting appropriate recovery goals, and/or setting appropriate estimates of maximum sustainable yield for management purposes.

This issue first obtained wide recognition by fisheries biologists, who noticed that what passed as "normal", or the reference baseline for certain populations, (in relation to distribution, population size, and other characteristics) tended to shift over generations due to the changing perceptions of wildlife biologists, as opposed to any change in the real baseline characteristics. They termed this trend the "shifting baseline syndrome", and it has become a major issue in conservation biology (Pauly 1995; Sheppard 1995).

Obtaining appropriate data on the "natural" status of a population (predating human effects) is very difficult. For example, all of the methods used to quantify such characteristics of contemporary populations, such as distance sampling or mark-recapture methods, cannot be used to infer changes in the past. However, here too genetic analyses can be helpful. Changes in population demography leave specific signatures in the DNA of individuals. Because the genetic characteristics of a population change relatively slowly, these signatures can be detectable for many generations. Applications have been developed to tease out many of these signatures, a few of which are discussed below and include: the 'M-ratio' test (Garza & Williamson 2001), heterozygosity excess (Cornuet & Luikart 1996), and Bayesian skyline-plots of population size change over time (Drummond 2005).

The 'M-ratio' test, which was developed by Garza and Williamson (2001) utilizes the formula $M=K/r$, which tests for a population bottleneck genetic signature, by examining the ratio of the number of microsatellite alleles (K) to the range in allele size (r) (Peery *et al.* 2012). When a population declines in size the number of alleles is

expected to be reduced faster than the range in allele size, because only when there is a loss in the largest and smallest allele will it reduce the range (*r*) and these allele are found in a lower frequency and therefore are less likely to be lost. This results in the ratio being smaller in a population that has recently declined (Garza & Williamson 2001; Peery *et al.* 2012). To test if this ratio is smaller than expected at equilibrium, the mean observed 'M-ratio' (across multiple loci) is compared to the calculated critical value, and if it is lower than a population bottleneck can be inferred (Garza & Williamson 2001).

In populations that have experienced a bottleneck event, expected heterozygosity calculated from the observed number of alleles will be lower than the actual observed heterozygosity (Cornuet & Luikart 1996; Peery *et al.* 2012; Torres-Florez *et al.* 2014). This is because when there is a loss of individuals the number of alleles is reduced faster than the heterozygosity, therefore the observed heterozygosity becomes larger than expected heterozygosity based on the number of alleles available in the sample set (Cornuet & Luikart 1996; Piry *et al.* 1999). This discrepancy between heterozygosity observed and expected is what genetic programs, such as BOTTLENECK, have been designed to detect (Piry *et al.* 1999).

The Bayesian skyline plot model is a method for estimating historical population demography from a sample set of sequences (Drummond 2005; Peery *et al.* 2012). Ho and Shapiro (2011) explained the skyline plot framework as estimating the genealogy of a population from the available sequence data and then estimating the demography based on the genealogy. The program BEAST is able to combine these two estimation into a single step (Drummond *et al.* 2002; Drummond 2005; Drummond & Rambaut 2007; Drummond *et al.* 2012). BEAST can take the sequences provided and estimate a

genealogy, or tree, based on the posterior probability of coalescing on the most recent common sequence. The program TRACER v1.6 (Rambaut & Drummond 2007) can take these trees constructed in BEAST and infer the changes in population size over time based on the branch length between coalescent events. The longer the branch length can infer reductions in population size because individuals, and their sequences, have been removed and therefore it will take longer to find a common sequence (Kuhner 2009). Similarly, shorter branch lengths between coalescent events can infer a population expansion because there is an increase in sequences leading to shorter times between coalescent events. Overall, this means that BEAST and TRACER are able to estimate the changes in population size from a given time point into the past. If the sequence data utilized are from contemporary samples, then the demography is from present day going backwards, but if ancient samples are used then the inferences can have a longer timeline (Stiller *et al.* 2010; Nyström *et al.* 2012). If the appropriate variables are accounted for, such as: mutation rate, substitution model, site model, MCMC chain length, then the skyline plots can prove to be very useful for helping to infer and understand large scale changes in population demography.

As an example of these approaches, Miller *et al.* (2012) examined the mtDNA of contemporary polar and brown bears to infer past demographic changes due to climate changes. They provided estimates of changes in effective population size over 5 million years and inferred that these changes may have been due to key historic climatic events (Miller *et al.* 2012). They suggest that the polar bear has gone through a dramatic and prolonged decrease in effective population size over the last 500,000 years as a result of environmental changes, indicating that this population may be particularly susceptible to

the direct and indirect effects of humans (Miller *et al.* 2012). This case study provides an example of how the analysis of contemporary samples can be used to infer historic patterns in the change in effective population size and attempt to match them to key changes in the environment.

A second approach where genetics can be used to improve our understanding of population history is when samples are available from pre-human specimens, and the genetic characteristics of these historic specimens can then be compared to those of contemporary populations to infer changes in population demography. For example, Nyström *et al.* (2012) examined DNA from woolly mammoths of the late Pleistocene. The researchers were able to separate the woolly mammoth samples into two time periods based on radiocarbon dating, and interestingly the two temporarily separated samples also differed in their amount of genetic variation (Nyström *et al.* 2012). Using genetic data from both mtDNA and microsatellites, it was inferred that the larger population once found in north-eastern Siberia went through a bottleneck event 12 thousand years BC and was left with a much smaller and less genetically varied population (Nyström *et al.* 2012; Amaral *et al.* 2012). Here, the researchers were able to compare sample sets that differed in age range and infer a change in demography that may have gone undetected if this age difference was unknown.

Whales are long-lived species with cryptic lifestyles for which it is very difficult to obtain accurate long-term population size estimates, and some species have largely unknown life histories. In recent years there has been a surge of interest in investigating pre-exploitation population sizes to better assess the status of populations prior to the onset of industrial whaling, and thus to better understand the effects of whaling on the

exploited species, as well as on the ecosystems as a whole (Rooney *et al.* 1999; 2001; Alter *et al.* 2007; 2012; Torres-Florez *et al.* 2014). Previous estimates of pre-exploitation population sizes were based on whaling ship's logbooks, barrels of whale oil obtained, and amount of baleen collected or exported (Smith & Reeves 2003; Carroll *et al.* 2014). There can be many issues with these methods, as logbooks may be incomplete, the species identification may be incorrect, variation in the amount of oil and/or baleen obtained from individuals, and they may not take into account the number of whales that were killed but not recovered. The use of genetic can help to clarify these population size estimates.

The field of inferring historical population demography is constantly changing and improving with the discovery and introduction of new and more powerful detection methods. These new techniques, be they molecular, statistical or analytical, can assist researchers in answering key historical questions about all whale species. Through the investigation of well-studied species, such as prairie chickens and how their genetics have responded to past bottleneck events, it is possible to use these techniques to infer the historical demography of whales, as their history is not well represented in the literature (Bellinger *et al.* 2003; Johnson *et al.* 2007).

Objectives

There are discrepancies between oral history, literature, and archaeological evidence as to which species were traditionally taken by First Nations groups, who live on the west coast of North America. Using genetic methods for species identification described above which utilize mitochondrial DNA, I hope to discover the whale species

that were traditionally preferred by the Nuu-chah-nulth First Nations and in what proportion. These findings will then be compared to the results obtained through previous research and hopefully put to rest any inconsistencies, or discover possible geographical differences in which species were hunted.

Secondly, the genetic data recovered from conducting species identification on these ancient whale samples will be used to infer the historic and pre-historic demography of those species that where decimated due to the commercial whaling industry. The importance of inferring historic and pre-historic demography is because of the interest of some First Nations groups to resume their traditional practice of hunting whales. If there is to be a resumption of traditional whaling, we need to ensure that these whale species have recovered and that the practice is sustainable. I will also attempt to recover nuclear DNA from these ancient samples to observe any changes in genetic variation due to the reductions in population sizes from commercial whaling. Additionally, using the genetic signatures found within the contemporary individuals, I will examine for signs of a population bottleneck, which we know occurred in the late 19^{th} and early 20^{th} century. Whale species are long-lived, with long generation times, and examining for signs of a known bottleneck event will check to see if enough time has passed since this reduction in numbers for it to be accurately reflected in the DNA. It will help to establish what the limitations are of these techniques of long-lived species.

Literature Cited

Allentoft ME, Collins M, Harker D *et al.* (2012) The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 4724–4733.

Alter SE, Newsome SD, Palumbi SR (2012) Pre-Whaling Genetic Diversity and Population Ecology in Eastern Pacific Gray Whales: Insights from Ancient DNA and Stable Isotopes (M Knapp, Ed,). *PLoS ONE*, **7**, e35039.

Alter SE, Rynes E, Palumbi SR (2007) DNA evidence for historic population size and past ecosystem impacts of gray whales. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 15162–15167.

Amaral AR, Beheregaray LB, Bilgmann K *et al.* (2012) Influences of past climatic changes on historical population structure and demography of a cosmopolitan marine predator, the common dolphin (genus Delphinus). *Molecular Ecology*, **21**, 4854–4871.

Bellinger MR, Johnson JJ, Toepfer J, Dunn P (2003) Loss of genetic variation in the greater prairie chickens following a population bottleneck in Wisconsin, USA. *Conservation Biology*, **3,** 717-724.

Bennett P (2000) Demystified ... microsatellites. *Molecular pathology*, **53**, 177–183.

Budowle B, Allard MW, Wilson MR and Chakraborty R (2003) Forensics and Mitochondrial DNA: Applications, Debates and Foundations. *Annual Review of Genomics and Human Genetic,s* **4,** 119-41.

Butler JM and Levin BC. (1998) Forensic applications of mitochondrial DNA. *Trends in Biotechnology,* **16,** 158-162.

Campbell NA, Reece JB (2008) *Biology*. 8th ed. Upper Saddle River (NJ): Benjamin Cummings. p.109-110

Carroll EL, Jackson JA, Paton D, Smith TD (2014) Two intense decades of 19th century whaling precipitated rapid decline of right whales around New Zealand and East Australia. *PLoS ONE*, **9**, e93789.

Cernetig M (1997) Natives set sail on whale of a controversy. *Toronto Globe and Mail*, 27 March

Clayton D (2000) *Islands of Truth: The Imperial Fashioning of Vancouver Island*. UBC Press, Vancouver.

Cooper A, Poinar HN (2000) Ancient DNA: do it right or not at all. *Science*, **289**, 1139–1139.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Coté C (2010) *Spirits of our Whaling Ancestors: Revitalizing Makah & Nuu-chah-nulth Traditions.* UBC Press, Vancouver.

de Bruyn M, Hoelzel AR, Carvalho GR, Hofreiter M (2011) Faunal histories from Holocene ancient DNA. *Trends in Ecology & Evolution*, **26**, 405–413.

Drummond AJ (2005) Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Molecular Biology and Evolution*, **22**, 1185–1192.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, **161**, 1307–1320.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.
Foote AD, Hofreiter M, Morin PA (2012) Ancient DNA from marine mammals: Studying long-lived species over ecological and evolutionary timescales. *Annals of Anatomy - Anatomischer Anzeiger*, **194**, 112–120.
Frasier TR, Koroscil SM, White BN, Darling JD (2011) Assessment of population substructure in relation to summer feeding ground use in the eastern North Pacific gray whale. *Endangered Species Research*, **14**, 39–48.
Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305–318.
Ho SYW, Shapiro B (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources*, **11**, 423–434.
Huelsbeck DR (1988) Whaling in the precontact economy of the central northwest coast. *Arctic Anthropology*, 1–15.
Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome b gene of mammals. *Journal of Molecular Evolution*, **32**, 128–144.
Johnson JA, Dunn PO, Bouzat JL (2007) Effects of recent population bottlenecks on reconstructing the demographic history of prairie-chickens. *Molecular Ecology*, **16**, 2203–2222.
Karp G (2005) *Cell and Molecular Biology: Concepts and Experiments.* 4th ed. John Wiley and Sons, Inc., Hoboken. 185 p.
Kuhner MK (2009) Coalescent genealogy samplers; windows into population history. *Trends in Ecology and Evolution*, **24**, 86-93.
Larizza A, Pesole G, Reyes A, Sbisa E, and Saccone C (2002) Lineage specificity of the evolutionary dynamics of the mtDNA D-loop region in rodents. *Journal of Molecular Evolution* **54:**145-155.
Linacre A, Tobe SS (2011) An overview to the investigative approach to species testing in wildlife forensic science. *Investigative Genetics*, **2**, 2.
Lindahl T (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.
McMillan AD, Claire DES (2005) *Ts'ishaa*. Archaeology Press. Department of Archaeology, Simon Fraser University., Burnaby.
Miller W, Schuster SC, Welch AJ *et al.* (2012) Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, E2382–90.
Monks GG, McMillan AD, St Claire DE (2001) Nuu-chah-nulth whaling: archaeological insights into antiquity, species preferences, and cultural importance. *Arctic Anthropology*, 60–81.
Mulligan CJ (2006) Anthropological applications of ancient DNA: Problems and prospects. *American Antiquity*, 365–380.
Nyström V, Humphrey J, Skoglund P *et al.* (2012) Microsatellite genotyping reveals end-Pleistocene decline in mammoth autosomal genetic variation. *Molecular Ecology*, **21**, 3391–3402.
Parson W, Pegoraro K, Niederstätter H, Föger M, Steinlechner M (2000) Species identification by means of the cytochrome b gene. *International Journal of Legal Medicine*, **114**, 23–28.
Pauly D (1995) Anecdotes and the shifting baseline syndrome of fisheries. *Trends in*

Ecology & Evolution, **10**, 430.
Peery MZ, KIRBY R, REID BN *et al.* (2012) Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular Ecology*, **21**, 3403–3418.
Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity*, **90**, 502–503.
Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, **6**, 847–846.
Pruvost M, Schwarz R, Correia VB *et al.* (2007) Freshly excavated fossil bones are best for amplification of ancient DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 739–744.
Reeves RR (2002) The origins and character of "aboriginal subsistence" whaling: a global review. *Mammal Review*, **32**, 71–106.
Rooney AP, Honeycutt RL, Derr JN (2001) Historical population size change of bowhead whales inferred from DNA sequence polymorphism data. *Evolution*, **55**, 1678–1685.
Rooney AP, Honeycutt RL, Davis SK, Derr JN (1999) Evaluating a putative bottleneck in a population of bowhead whales from patterns of microsatellite diversity and genetic disequilibria. *Journal of Molecular Evolution*, **49**, 682–690.
Sheppard C (1995) The shifting baseline syndrome. *Marine Pollution Bulletin*, **30**, 766–767.
Smith TD, Reeves RR (2003) Estimating American 19th Century Catches of Humpback Whales in the West Indies and Cape Verde Islands. *Caribbean Journal of Science*, **39**, 286–297.
Stiller M, Baryshnikov G, Bocherens H *et al.* (2010) Withering away--25,000 years of genetic decline preceded cave bear extinction. *Molecular Biology and Evolution*, **27**, 975–978.
Tajima F (1996) The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. *Genetics*, **143**, 1457–1465.
Tobe SS, Kitchener AC, Linacre AMT (2010) Reconstructing mammalian phylogenies: A detailed comparison of the cytochrome b and cytochrome oxidase subunit I mitochondrial genes. *PLoS ONE*, **5**, e14156.
Torres-Florez JP, Hucke-Gaete R, Rosenbaum H, Figueroa CC (2014) High genetic diversity in a small population: the case of Chilean blue whales. *Ecology and Evolution*, **4**, 1398–1412.
Veeramah KR & Hammer MF (2014) The impact of whole-genome sequencing on the reconstruction of human population history. *Nature Reviews Genetics* **15**, 149–162
Verhovek SH (1999a) After the hunt, bitter protest and salty blubber. *New York Times*, 19 May.
Verhovek SH (1999b) Reviving tradition, tribe kills a whale. *New York Times*, 18 May.
Weber JL, Wong C (1993) Mutation of human short tandem repeats. *Human Molecular genetics*, **2**, 1123–1128.
Yang DY, Speller CF (2006) Co-amplification of cytochrome b and D-loop mtDNA fragments for the identification of degraded DNA samples. *Molecular Ecology Notes*, **6**, 605–608.
Yang DY, Watt K (2005) Contamination controls when preparing archaeological remains for ancient DNA analysis. *Journal of Archaeological Science*, **32**, 331–336.

Chapter 2

Genetic Species Identification of Bones from Traditional Whaling Sites on Vancouver Island

Abstract

Identifying what species were traditionally hunted by First Nations whaling communities is important from both a cultural and a biological perspective. Conflicting versions of the oral history of First Nations traditional whaling practices have been partially clarified with the help of anthropology and archaeology, but now the field of genetics can also be of assistance. This is of particular importance given the wishes of some groups to resume their traditional whaling practices, where the discovery of which species were targeted is fundamental. Species identification of historical bone specimens can be problematic, particularly when based on osteological characteristics, which can result in incorrect identifications and may be biased towards only identifying species for those skeletal elements that are recovered intact. Genetic methods, such as using the cytochrome-*b* gene and control region of the mitochondrial DNA (mtDNA), have proven to be more reliable and are also not limited to fully intact specimens. Using these target regions and conducting phylogentic analyses of the sequences, I was able to show that Nuu-chah-nutlh First Nations members from the central Vancouver Island hunted primarily gray whales, however humpback whales were also taken in large proportions. This information can assist in updating and accurately communicating Nuu-chah-nulth history to future generations, provide a better understanding about the historical exploitation of whale populations before commercial whaling, and finally to help make informed decisions regarding the potential resumption of traditional First Nations whaling practices.

Introduction

A few whale species that were once decimated by whaling through the early 1900s have increased to the point where they have arguably "recovered", with some being down-listed from Endangered Species Lists in North America. For example, eastern North Pacific gray whales (*Eschrichtius robustus*) were removed from the US Endangered species List in 1994, and the North Pacific humpback whale (*Megaptera novaeangliae*) was recently downgraded from "threatened" to "special concern" under Canada's Species At Risk Act (COSEWIC 2011). With this recovery has come an increase in interest to resume traditional whaling practices by some First Nations groups, and some have negotiated treaties that will allow the resumption of whaling (Coté 2010).

With the potential for these renewed whaling practices comes the responsibility to ensure proper population management from a biological perspective, and to ensure

historical accuracy from a cultural perspective. From a cultural perspective, if the proposed hunts are promoted as a continuation of important historical and cultural practices, which have not been conducted for several generations, it is important to ensure they accurately reflect historic hunts, through the synthesis of oral history and archaeological evidence. From a biological perspective, it is important that relevant management plans are based on appropriate scientific data to ensure sustainability. Meeting these responsibilities requires accurate information on which species were the target of sustenance whaling, and in what proportions. Such information will ensure that contemporary practices are consistent with those of the past and help recover information that was lost due to the gap in whaling practices for multiple generations.

The Nuu-chah-nulth on British Columbia's Pacific Northwest is an example of one such First Nations group that may resume whaling. The Nuu-chah-nulth is made up of 14 smaller communities found along the west coast of Vancouver Island and they have a long presence in the area, which can be seen through an abundance of oral history and archaeological artifacts (Monks *et al.* 2001; Coté 2010). Traditional whaling practices are intertwined throughout the cultural history and have been dated back thousands of years (Harkin 1998; Monks *et al.* 2001; McMillan & Claire 2005; Coté 2010).

Although this traditional whaling took place for thousands of years, it was the onset of industrial commercial whaling that decimated local populations, at which point the Nuu-chah-nulth voluntarily stopped their whaling activities (Reeves 2002; Coté 2010). This stoppage took place in the early 1900s, which was prior to the moratorium on commercial whaling that was implemented by the International Whaling Commission (IWC) in 1986 (Reeves 2002; Coté 2010). The loss of whales not only affected the

Nuu-chah-nulth, but also other First Nations cultures along the west coast of Canada, such as the Haida, for whom whaling was also important (Acheson & Wigen 2002).

Some First Nations groups have proposed that this lack of traditional sustenance whaling was not intended to be permanent. The Makah, a whaling First Nations group on the west coast of Washington State, have been in negotiations and litigation since the mid 1990s to resume their traditional whaling practices of the gray whale (Greenlee & Dunnell 2010; Coté 2010; Frasier *et al.* 2011) With the current legal efforts and media coverage surrounding this case, there is potential for the outcome to influence Canadian perspectives and negotiations on the resumption of traditional whaling. In fact, a recent treaty agreement between the Maa-nulth, who are a subgroup of the Nuu-chah-nulth First Nation, the Province of British Columbia, and Government of Canada came into effect in April of 2011 (Coté 2010). Included in this newly negotiated treaty was their right to take whales, including, but not limited to, gray and sei whales, although a side-agreement states that they will not include these two species in their annual fishing plan for 25 years from the 2011effective date (Indian and Northern Affairs Canada 2006; Coté 2010). Whaling is a main focus within Nuu-chah-nulth oral history, and many believe that sustenance whaling is key to maintaining the culture.

The Nuu-chah-nulth oral history, and the archaeological evidence found on the west coast of Vancouver Island, suggest that their whaling practices date as far back as 3500 years ago (Monks *et al.* 2001; Monks 2003; Coté 2010). One major question of these historic practices is which whale species were targeted. Osteological analysis of whale bones recovered from these historic sites resulted in the identification of many species, with gray and humpback whales being the most prevalent (Huelsbeck & Fisken

1983; Huelsbeck 1988; Monks *et al.* 2001; McMillan & Claire 2005). However, two issues arise when using osteological characteristics for species identification. First, most bones recovered are fragmented, and therefore not suitable for species identification, resulting in species being identified for only a limited portion. For example, at one site fewer than 50% of all bones recovered were identified to species because most of the bones were too fragmented for analysis or there was a lack of comparative material (Huelsbeck & Fisken 1983; Huelsbeck 1988). If intact bones are required for osteological species identification, then bones that are fragmented or were later modified into tools can be unsuitable for analysis (Monks *et al.* 2001; Newman *et al.* 2002; McLeod *et al.* 2008; Greenlee & Dunnell 2010). This bias in selecting samples large and intact enough for osteological analysis could lead to a bias in the ratio of the estimated importance of different species.

Second, osteological species identification from historic whale bones has a tenuous history, with several cases where such identification was subsequently found to be incorrect. For example, Cumbaa (1986) used osteological characteristics to identify the species of 17 humeri recovered from a 16th century Basque whaling site, concluding that 9 originated from bowhead whales (*Balaena mysticetus*) and 8 from North Atlantic right whales (*Eubalaena glacialis*). This 50:50 ratio was subsequently used to estimate pre-exploitation population size for the right whale, and to set recovery goals. However, subsequent genetic analyses of these same bones showed that 16 were from bowhead whales, and only one from a right whale (Rastogi *et al*. 2004). Further analyses of a much larger sample set confirmed the abundance of bowhead whales, and lack of right whales, in these hunts (McLeod *et al*. 2008). Because so much weight was placed on the original

50:50 ratio, the genetic data showing these data were incorrect had large implications for the understanding of the history, current status, and recovery potential of these species.

One solution to both of these issues is using genetic analyses for species identification. First, genetic analyses require only miniscule amounts of bone, and therefore can be conducted as easily on small fragments as on large intact bones. This property greatly increases the potential sample sizes for analyses of historic whale bones, and eliminates any potential biases caused by the requirement of intact bones. Second, species identifications based on genetic analyses are extremely reliable, particularly for such well-studies species as the large baleen whales. Indeed, genetic analyses have become the primary means of species identification in whales, particularly when only small samples are available (Baker & Palumbi 1994; Dalebout *et al.* 2002).

Given the cultural and biological importance of accurately identifying which species were targeted in traditional hunts, I conducted genetic analyses of whale bones collected from a variety of historic Nuu-chah-nulth whaling sites from coastal Vancouver Island. By using molecular species identification methods, it is possible to accurately determine which species were hunted and in what proportion. This could eliminate the bias and issues associated with osteological identification and provide an accurate range of the species taken by these groups. Comparing these data to those from other areas along the North American west coast also provides a means to assess if there were geographical differences in which species were targeted.

Materials and methods

Site selection and sample collection

The District of Tofino, on the west coast of Vancouver Island, is located within

traditional Nuu-chah-nulth First Nations territory. The Tla-o-qui-aht are members of the Nuu-chah-nulth First Nation that reside in the Tofino area and they once practiced sustenance whaling. Potential sampling locations were identified through speaking with members of this First Nations group and examining the history of the area (Clayton 2000). One identified key site was Echachist Island, which was a summer settlement and was used as a whale butchering site (Clayton 2000; Joe Martin, personal communication July 2013). Bones were collected from 9 known traditional sites and 1 commercial whaling station (in operation form 1905-1918 by The Pacific Whaling Company) in and around Tofino during the summer months of 2011-2013 (Nichol *et al*. 2002) (Figure 2.1 & Table 2.1). Bones were recovered from three types of locations: (a) the surface, on shore or inland near residential areas; (b) just off shore in low tide; or (c) from test digs on shell middens. Bone shavings for species identification were collected from the larger bones or the bone fragments themselves were collected. The bone samples were scored on a 5 point scale for fragmentation, with 5 meaning the bone was completely intact with only minor chipping to the bone, and 1 representing a bone that was completely fragmented and unidentifiable (Table 2.2). Those bone samples with a fragmentation score of 3 or less were considered not identifiable based solely on osteological characteristics due to the lack of any distinguishable characteristics.

Sampling at the historical whaling sites was conducted under the permission of Tla-o-qui-aht members, and those samples that were recovered from what is now private property was done with the permission of the land owners.

Radiocarbon Dating

As part of a related ongoing project, radiocarbon dating was conducted on 6 of the bones. The purpose of dating these samples was to provide a rough time frame for bones in the area, and to demonstrate that those found on the surface of the sampling sites could in fact be classified as ancient. The 6 bones that were selected were those samples with enough shavings (1-2 grams) available without compromising the amount necessary for species identification. Five of the bones were from the main sampling site of Echachist Island and the sixth was from Hesquiat Island (Figure 2.1).

DNA extraction

DNA was extracted from 0.15-0.18 grams of bone shavings using a modified Qiagen column extraction as per Rastogi *et al.* (2004) and McLeod *et al*. (2008). Whaling traditions on the West coast of Vancouver Island date back thousands of years, therefore the DNA from recovered bones was likely of low quality and quantity, and highly susceptible to contamination from contemporary samples (Mulligan 2006). Thus, special protocols needed to be followed when working with ancient DNA (aDNA). As suggested by Cooper and Poinar (2000), all aDNA processing, prior to PCR amplification, was carried out in a dedicated and isolated laboratory where analysis on extant cetaceans has never been preformed, and negative controls were included at each step of extraction and amplification to ensure there was no contamination (Yang & Watt 2005).

mtDNA Amplification and Qualification:

For each sample, the control region and cytochrome-*b* gene of the mtDNA were

targeted for amplification. A ~450 base pair (bp) region of the control region was amplified using the primers t-PRO and Primer-2 from Yoshida *et al.* (2001). A ~500 bp region of the cytochrome-*b* gene was amplified using the primers CBMYSTF1 and CBMYSTR from McLeod *et al.* (2008). PCR amplifications were conducted in 20µl volumes containing: 5µL DNA extract, 1X PCR buffer (Promega), 1.5mM $MgCl_2$, 0.2mM each dNTP, 0.3µg/µL bovine serum albumin (BSA), 0.1U/µL *Taq* polymerase (Promega), and 0.3µM of each primer. PCR cycling conditions for the cytochrome-*b* gene consisted of an initial 5 minute denaturation step at 94°C; 50 cycles of 94°C for 30 seconds, 60°C for one minute, and 72°C for one minute; with a final extension step at 60°C for 45 minutes. The PCR cycling conditions for the amplification of the control region were different, as they were conducted using touchdown PCR. Touchdown PCR consisted of an initial 5 minute denaturation step at 94°C; 50 cycles of 94°C for 30 seconds, 65-55°C for one minute (see below), and 72°C for one minute; with a final extension step at 60°C for 45 minutes. The starting annealing temperature was 65°C, which was lowered by 0.5°C every cycle until reaching 55°C, where it remained for 30 cycles.

To check the quantity and quality of the amplification process, PCR product for each sample was run on 1.5% agarose gels stained with 0.5µg/mL ethidium bromide and visualized under UV light. In addition to a Low Mass DNA Ladder (Invitrogen), amplification standards made from a contemporary gray whale sample and at the raw DNA concentrations of 1ng/µl, 0.1ng/µl and 0.01ng/µl, were amplified in each reaction and visualized on the agarose gel. This was done to confirm that the samples were amplifying at the targeted region, to check for possible contamination, and to assist in

estimating PCR product concentrations of each sample, because the sequencing procedure requires 5ng of DNA per 100 base pairs of the target region, per reaction.

DNA Sequencing:

In preparation for sequencing, the PCR product was purified to remove excess dNTPs and primers that were not incorporated during the initial amplification. This process was conducted in an enzymatic reaction using reagents from New England BioLabs®, which required 0.65ul Antarctic Phosphatase Buffer, 0.1µl Antarctic Phosphatase, and 0.03µl Exonuclease I per 5µl sample of PCR product. The thermal cycling conditions were 15 minutes at 37°C, followed by 15 minutes at 80°C and finally held at 10°C.

Sequencing was performed on both target regions and in both directions using the Sanger *et al*. (1977) method. The sequencing PCR cocktail, with all reagents from Applied Biosystems®, took place within a 15µl reaction and contained 0.25X Big Dye Terminator reaction mix, 1X Big Dye Terminator sequencing buffer, 1µl/reaction of 5µM of specified primer, and 5µl/reaction of 4.5ng/µl (for control region) or 5ng/µl (for cyyochrome-b) DNA. The PCR cycling consisted of an initial 2 minute denaturation step at 96°C; 30 cycles of 96°C for 20 seconds, 50°C for 20 seconds, and 60°C for 4 minutes; with a final hold step at 4°C.

Once the sequencing reaction was completed, an ethanol precipitation procedure was conducted to de-salt the samples (Irwin *et al.* 2003). The resulting DNA was re-suspend in 10µl HiDi formamide and analyzed on an ABI 3500xL genetic analyzer for sequencing.

Sequence Analysis:

Using the 4Peaks software Version 1.7.2. (Griekspoor and Groothuis 2006), all successful sequences were edited by eye to ensure correct base assignment.

The forward and reverse compliment sequences for each sample were aligned in ClustalX (Thompson *et al.* 1997; Larkin *et al*. 2007). The alignment showed where the two sequences could be combined into one consensus sequence. This was completed for both target regions for all samples. Therefore, each sample had a consensus sequence for the cytochrome-*b*, control region, or both.

All available consensus sequences were compared to cytochrome-*b* and control region sequences available on the GenBank database using the Basic Local Alignment Search Tool (BLAST) algorithm for initial species identification (Altschul *et al.* 1990; 1997; Benson *et al.* 2013). The results of interest were those which had the maximum-scoring segment pair (MSP), that is, the best match to the query sequence (Altschul *et al*. 1990; Parson *et al*. 2000). For those bone samples that had a consensus sequences available at both target regions, each was compared to the database to ensure consistent species identification.

Phylogenetic Analysis:

Phylogenetic analysis was then conducted to confirm preliminary BLAST species identification. This was done because the highest sequence match in BLAST may not be the closest relative phylogenetically (Koski and Golding 2001). The highest match may also be to a sequence submitted by an independent laboratory and not from a published,

peer-reviewed article. This means that sequences in the database may contain errors, resulting in incorrect species identification (Harris 2003).

Reference sequences, for the cytochrome-*b* and control region, were found on GenBank for 14 species within the suborder Mysticeti (Table 2.3). Sequences were aligned with ClustalX and checked by eye using a range of gap opening and extension penalties, to ensure that alignments correctly accounted for such changes as additions or deletions in the sequences. The program ModelGenerator v85 (Keane *et al*. 2006) was used to identify the most appropriate model of molecular evolution and then TREE-PUZZLE (Schmidt *et al*. 2002) was used to estimate the transition:transvertion ratio, as well as the α value describing the shape of the gamma distribution if there was found to be rate heterogeneity in substitution rates across sites.

A Bayesian approach was used for the phylogenetic analysis, as implemented in the program MrBayes (Huelsenbeck & Ronquist 2001; Ronquist *et al.* 2012). The MrBayes program is ideal because it allows genetic information from different regions to be combined into a single analysis, while still allowing for different substitution models for each region. This allows users to combine multiple regions for analyses without the known issues that arise from simply concatenating regions for analysis (e.g., Kubatko and Degnan 2007). Thus, species identification results were based on the combined information from both the cytochrome-*b* gene and control region. Two MCMC runs were conducted, and then the resulting phylogenetic trees were summarized, with 6 chains and with a burn-in of 20,000 generations each. The analysis was run for 5,000,000 generations, but a stop value of 0.01 for the average standard deviation of split frequencies was implemented to ensure the program had run long enough to obtain

accurate results. The variables accounted for in the phylogenetic analysis were incorporated and the most likely model of molecular evolution, from ModelGenerator, was HKY+G for both mtDNA regions and with alpha values of 0.20 and 0.71 for the cty-b and control regions, respectively. The transition:transversion ratio, which was estimated using TREE-PUZZLE, was 15.76 for the cytochrome-*b* region and 4.95 for the control region.

Minimum Number of Individuals

Haplotype and nucleotide differences between the sequences were identified using Fabox v1.4 (Huelsbeck & Fisken 1983; Huelsbeck 1988; Baker & Palumbi 1994; Dalebout *et al.* 2002; Villesen 2007). Unique cytochrome-*b* haplotypes and unique control region haplotypes were identified, and combined to result in a series of unique haplotype combinations that were used to estimate the minimum number of individuals (MNI) sampled. Bones that did not have a unique haplotype combination may be from the same individual or may just have the same haplotype, as mitochondrial DNA is maternally inherited, and therefore individual identity cannot not be distinguish based on mitochondrial markers alone. Additionally, samples that possessed the same haplotype combination, but were recovered from different sample sites, such as different islands, were assumed to be different individuals, as bones were not valued in west coast whaling societies and usually left on the beach at the site of butchering (Monks et al. 2001).

Results

One hundred and nine bone samples were collected from the 10 sites around the Village of Tofino (Figure 2.1 and Table 2.1). Radiocarbon dates of the 6 dated bones ranged from 499-1139 years before present (BP=1950). This dating has been corrected for the Marine Reservoir Effect for the area (McNeely *et al*. 2006). The age range of the bones pre-date commercial whaling, and therefore can be classified as ancient for the purpose of this study.

Of the 109 samples collected, 68 were sequenced at both the cytochrome-*b* gene and the control region. Seventeen only sequenced at the control region and 4 only sequenced at the cytochrome-*b* gene region. These 89 samples were all identified to species and the identifications were consistent for those samples that had sequences from both regions available (Table 2.4).

One hundred and nine samples, with the possibility of being sequenced at two regions, gives a total of 218 possible sequences to identify to species and 157 sequences were recovered from the two target regions. This is a ~72% success rate of recovering mitochondrial DNA from these ancient whale bone samples. This sequencing success rate is high when compared to other studies involving ancient DNA extraction and sequencing, which range from 40-60% (Nichol *et al.* 2002; McLeod *et al.* 2008; Gravlund *et al.* 2012; Grier *et al.* 2013). This high success rate may be due to the age of the bones compared to those studies that had a lower success rate, as the bones in this study may be younger and therefore the DNA will not be as degraded.

Eighty-six of the bones were fragmented enough that species identification would not likely be possible based on osteological characteristics. However, genetic analysis of

these bone fragments was very successful (Table 2.2). Particularly for Echachist Island, where 63 of 68 bones that were scored as </=3 were identified to species. One site that was not as successful was Sechart Whaling Station, as only 1 of 7 bones scored as </=3 was identified (Table 2.2). This low success rate at Sechart Whaling Station may have been due to the method in which commercial whaling companies processed the bones in order to extract as much oil as possible, such as boiling them, but further investigation into the specific practices at this station are necessary.

Of the 89 sequenced bones, species identification from the BLAST analyses identified 42 as gray whale, 37 as humpback whale, 7 as North Pacific right whales (*Eubalaena japonica*), and 3 as fin whale (*Balaenoptera physalus*) (Table 2.4). The phylogenetic analysis with MrBayes resulted in the same species identifications as the BLAST analysis (Table 2.4 and Figure 2.2).

The above results are for each bone, without trying to correct for multiple bones that may represent the same individual. When haplotype definitions and unique haplotype combinations were assigned using FaBox, the minimum number of individuals (MNI) identified were: 26 gray whale, 16 humpback whale, 4 North Pacific right whale, and 2 fin whale (Table 2.4). As expected, based on the properties of the cytochrome-*b* gene and the control region of mitochondrial DNA, there were a higher number of control region haplotypes within each species and this increased the resolution of the MNI estimates (Table 2.5).

Discussion

Understanding Whaling Practices

Investigating the oral history of First Nations whaling on the west coast of North America can prove quite difficult because the literature can give multiple, and at times conflicting, versions of the same event. This has proved to be true when examining the literature to determine which species of whale was the primary and preferred target of traditional sustenance whaling (Kool 1982; Coté 2010). There is still a discrepancy, on which species was preferred, among the Tla-o-qui-aht First Nations. Some members believe that humpback whales were the preferred targets, similar to neighboring whaling groups, though some of the literature points towards gray whales (Kool 1982; Monks *et al*. 2001). In the past few decades, archaeological evidence attempted to clarify this confusion, but was limited to osteological analysis for species identification and in the end did not help as much as anticipated, as a large majority of the bones were too fragmented to be identified to a single species (Monks *et al.* 2001). There are those that believe, based on oral history in the area and archaeological evidence, that humpback whales were actively targeted, similar to the Tseshaht First Nations of Barkley Sound. This is interesting because it differs from the Makah of Washington State who primarily sought gray whales, though this was concluded based on osetological species identification (Kool 1982; Huelsbeck & Fisken 1983; McMillan & Claire 2005; Coté 2010). There is a lack of resolution when using traditional morphological comparative methods on their own, but genetics methods can assist in confirming species identification (Yang *et al*. 2005; Grier *et al*. 2013). Additionally, genetic species identification can be performed on animal remains other than bone, such as baleen (Sinding et al. 2012). It was found in this study that bones, even those fragmented beyond skeletal element identification, could be identified to a species.

When I examined the proportion of species identified for the largest sample site of Echachist Island, which has a long history for the Tla-o-qui-aht First Nations as a summer dwelling, and the butchering of hunted whales usually took place on the nearby beach (Clayton 2000), the MNI of gray and humpback recovered and identified were 20 and 12, respectively. This is only a preliminary sampling of the vast quantity of bones available at this site, as some bones were buried or just offshore in too deep of waters to collect, but demonstrates that more gray whales were taken or were indeed the preferred species. These findings conflict with the archaeological evidence of the nearby Barkley Sound, but again, this could be due to the differing methods of species identification (Cumbaa 1986; Rastogi *et al.* 2004; McLeod *et al.* 2008). Additionally, this could imply that there may have been a difference of whale targets based on the geographical location of the First Nations groups along the west coast of North America and that researchers cannot make inferences on the preference of one group based on the evidence available from other (Monks et al. 2001). Further investigation of larger samples size from these, and other, whaling sites would clarify this issue. It would also be useful to obtain radiocarbon dates for more of the bone samples recovered and from different historical whaling sites, as there may have not only been a geographical difference in whale preference but a shift in target species over time depending on which species was found in greater numbers during the summer hunting months.

What This Means for the Whales

In regards to the proposition of resuming the practice of sustenance whaling, uncovering whether or not these groups had a preference for a particular species is now

a key priority. The treaty that was recently put into effect between the Province of British Columbia, Canada and the Maa-nulth First Nations does not specify a species that can be taken, though the side agreement to the treaty states that gray and sei whales cannot be taken until the year 2036 (Indian and Northern Affair Canada 2006). It even specifically states, "…sei whales have recovered from industrial exploitation and are no longer considered by Canada to be endangered species." This line is of particular concern as, according to COSEWIC, the sei whales found in the Pacific Ocean off the west coast of Vancouver Island are still considered endangered and their status was just recently evaluated in the spring of 2013 (COSEWIC 2013). Indeed, none have been seen, despite extensive effort, for many years, raising the possibility that they have become locally extirpated. Additionally, in the course of this research, sei whales were not mentioned in either the oral history or literature as a preferred target species, nor where they identified at any of the 10 sample sites investigated. How can such outlines be negotiated into a treaty agreement, when there is clearly a lack of supporting evidence that this species has recovered from previous exploitation or was even a traditional whaling target? Science and its research should be strongly considered in policy and treaty development, as well as litigation, when it involves the management of species that are endangered, threatened, or have very recently recovered, to ensure their survival. The fact that this treaty has already been agreed upon by the Maa-nulth, and the provincial and federal governments, means that the deadline for ensuring that the correct species were a part of these negotiations has passed. However, hopefully this—and similar—research can be incorporated into future treaties and management plans. Cetaceans have long life spans and generation times, so delaying the hunting of whales of the west coast off Vancouver

Island for 25 years, may not be adequate time to make sure these species are still increasing in population size, or in the case of the sei whale, to ensure that it will have enough time to recover from its endangered status (Baker *et al*. 1993; Taylor *et al*. 2007). Furthermore, if both gray and humpback whales were targeted equally, this could mean that when management decisions are made concerning future catch limits of whales, there would be allocation between the species and therefor reduce the overall hunting pressure on just one.

The discovery of which species was the historical target for the Nuu-chah-nulth First Nations will not only be important for the management of the whales, but will also be beneficial for maintaining and accurately sharing this cultural history to future generations. In recent years there has been a revitalization of First Nations traditions and a wish for them to be more prominent in Canadian culture. The investigation here will help shed light onto the long history found on the west coast and ensure it is based on the best evidence available.

Moving Forward

The information gathered from analyzing the genetic data will have many applications in addition to investigating the cultural aspects of First Nations traditional whaling. It will also be beneficial to incorporate these data into current demographic models of cetaceans found in the east Pacific Ocean and assist in establishing accurate accounts of the history of these whale species prior to their mass depletion, due to the commercial whaling industry (Rooney *et al*. 2001; Foote *et al*. 2012; Alter *et al*. 2012). It could give researchers a more accurate baseline for these species and help infer what

changes in demography may have taken place well before the onset of commercial whaling, and to what effect traditional sustenance whaling had on the population size and genetic variability of Pacific cetaceans (Alter *et al*. 2012; Torres-Florez *et al*. 2014). Finally, having an accurate population baseline for these species will be necessary to ensure that their recovery goals are accurately met and, if whaling does resume on Canada's west coast, that it is done in a responsible manner to ensure the survival of these species.

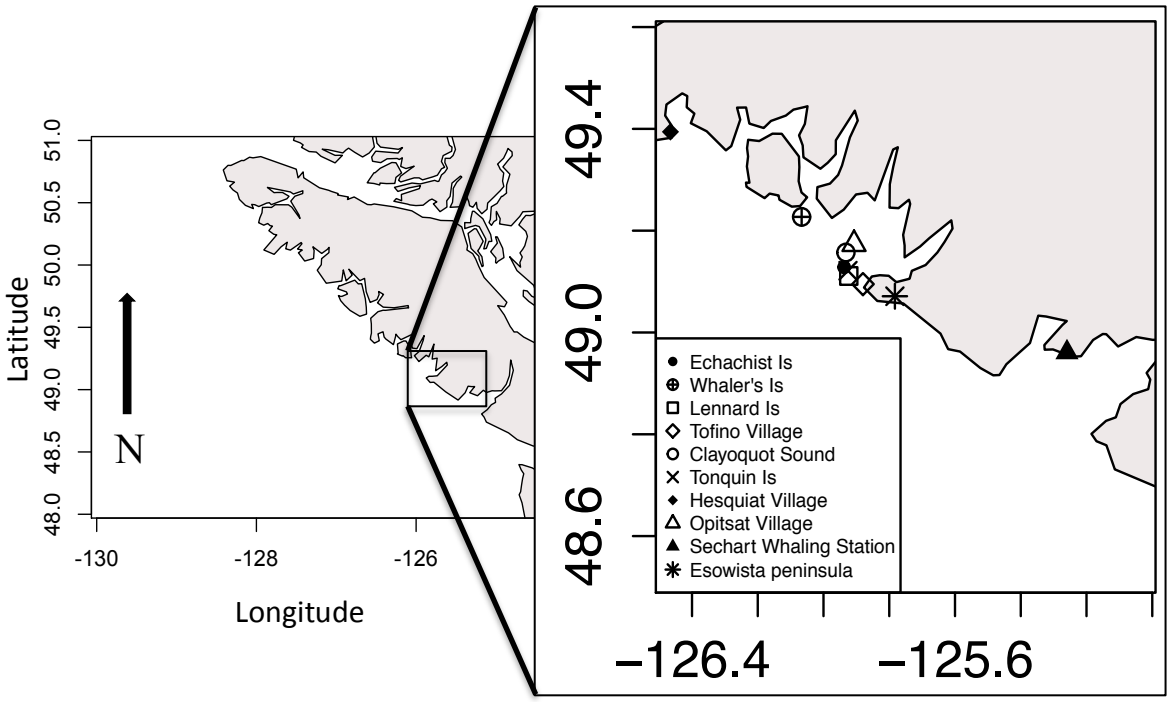

Figure 2.1. 10 sample sites off the west coast of Vancouver Island in British Columbia. Nine sampling sites are traditional Nuu-chah-nulth First Nations sites, while the tenth is an abandoned commercial whaling station.

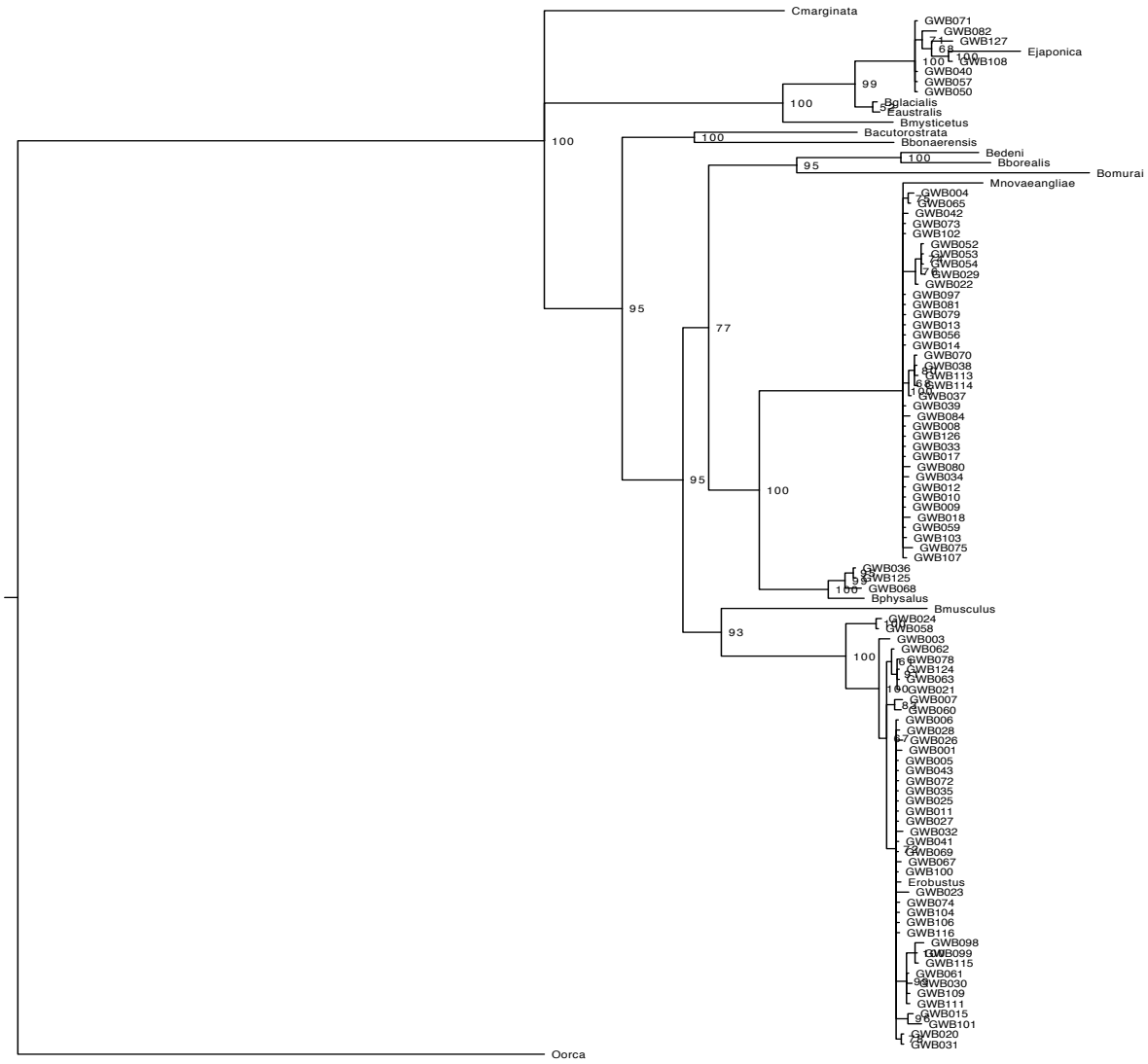

Figure 2.2. Resulting phylogenetic tree for species identification from MrBayes analysis. The numbers located on the branch nodes is the posterior probability (Huelsenbeck *et al.* 2001).

Table 2.1. GPS coordinates of ancient bone sampling locations around West Vancouver Island.

Sample Site	GPS Coordinates (decimal degrees)	# Samples Collected	# Samples Identified
Clayoquot Sound	49.157124, -125.931816	1	1
Esowista Peninsula	49.071167, -125.782842	1	1
Lennard Island	49.110961, -125.923076	2	2
Whaler's Island	49.227071, -126.066499	1	1
Echachist Island	49.128277, -125.937403	79	72
Tonquin Island	49.121326, -125.925508	1	1
Village of Tofino	49.153980, -125.902305	1	1
Sechart Whaling Station	48.9608, -125.2594	11	1
Opitsat Village	49.173504, -125.907170	4	3
Hesquiat Village	49.394220, -126.465575	8	6

Table 2.2. Number of samples with a fragmentation score* of 3 or less for 109 bone samples collected from 10 sites and the number that were identified to species using genetic methods.

Sample Site	No. of bones with a score of =/< 3	No. identified to species
Clayoquot Sound	0	N/A
Esowista Peninsula	0	N/A
Lennard Island	0	N/A
Whaler's Island	0	N/A
Echachist Island	68	63
Tonquin Island	1	1
Village of Tofino	0	N/A
Sechart Whaling Station	7	1
Opitsat Village	3	2
Hesquiat Village	7	5

*5: Completely intact, slight chipping; 4: Slight fragmentation, some skeletal elements missing; 3: Fragmentation, can still identify to bone type i.e. vertebrae, rib, etc. but species ID may be difficult; 2: Major fragmentation, difficulty identifying to bone type; 1: Complete fragmentation, cannot identify to bone type.
N/A = Not Applicable

Table 2.3. Reference sequences, for the cytochrome-*b* and control region, were found on GenBank for 14 species within the suborder Mysticeti and *O.orca* as the outgroup.

Species	Target Region	Accession Number	Reference
E. glacialis (North Atlantic right whale)	Cytochrome-*b*	X75587.1	Arnason and Gullberg 1994
	Control Region	X72199.1	Arnason *et al.* 1993
E. japonica (North Pacific right whale)	Cytochrome-*b*	AP006474.1	Sasaki *et al*. 2005
	Control Region	AP006474.1	Sasaki *et al.* 2005
E. australis (southern right whale)	Cytochrome-*b*	DQ095153.1	Kaliszewska *et al.* 2005
	Control Region	AF395044.1	Malik *et al.* 2000
B. mysticetus (bowhead whale)	Cytochrome-*b*	JF489130.1	Naidu *et al.* 2012
	Control Region	X72197.1	Arnason *et al.* 1993
C. marginata (pygmy right whale)	Cytochrome-*b*	X75586.1	Arnason and Gullberg 1994
	Control Region	X72198.1	Arnason *et al.* 1993
E. robustus (gray whale)	Cytochrome-*b*	X75585.1	Arnason and Gullberg 1994
	Control Region	X72200.1	Arnason *et al.* 1993
M. novaeangliae (humpback whale)	Cytochrome-*b*	X75584.1	Arnason and Gullberg 1994
	Control Region	X72202.1	Arnason *et al.* 1993
B. acutorostrata (minke whale)	Cytochrome-*b*	X75753.1	Arnason and Gullberg 1994
	Control Region	X72006.1	Arnason and Gullberg 1994
B. bonaerensis (Antarctic mike whale)	Cytochrome *b*	X75581.1	Arnason and Gullberg 1994
	Control Region	EF113823.1	Pastene *et al.* 2007
B. edeni (pygmy Bryde's whale)	Cytochrome-*b*	X75583.1	Arnason and Gullberg 1994
	Control Region	X72196.1	Arnason *et al.* 1993
B. omurai (Omura's whale)	Cytochrome-*b*	AB201256.1	Sasaki *et al.* 2006
	Control Region	AB201256.1	Sasaki *et al.* 2006
B. borealis (sei whale)	Cytochrome-*b*	X75582.1	Arnason and Gullberg 1994
	Control Region	X72195.1	Arnason *et al.* 1993
B. physalus (fin whale)	Cytochrome-*b*	EU303337.1	McLeod *et al.* 2008
	Control Region	FJ832130.1	Caputo and Giovannotti 2009
B. musculus (blue whale)	Cytochrome-*b*	EU303340.1	McLeod *et al.* 2008
	Control Region	X72204.1	Arnason *et al.* 1993
O. orca (killer whale)	Cytochrome-*b*	HQ405752.1	Foote *et al*. 2011
	Control Region	HQ405752.1	Foote *et al*. 2011

Table 2.4. Species breakdown and Minimum Number of Individuals (MNI) in the sample set based on unique haplotypes and sampling location using BLAST and MrBayes.

Species	**Total # of species identified**	**MNI**
Eschrichtius robustus	42	26
Megatera novaengliae	37	16
Eubalaena japonica	7	4
Balaenoptera physalus	3	2

Table 2.5. Number of unique cytochrome-*b* gene and control region haplotypes within the 4 identified species.

Species	**Cytochrome-*b* gene Haplotypes**	**Sequence Length (bp)**	**Control Region Haplotypes**	**Sequence Length (bp)**
Eschrichtius robustus	6	535	23	366
Megatera novaengliae	4	477	10	353
Eubalaena japonica	2	539	4	352
Balaenoptera physalus	2	540	2	431

Literature Cited

Acheson S and Wigen R (2002) Evidence for a prehistoric whaling tradition among the Haida. *Journal of Northwest Anthropology,* **36**, 115-168

Alter SE, Newsome SD, Palumbi SR (2012) Pre-whaling genetic diversity and population ecology in eastern pacific gray whales: Insights from ancient dna and stable isotopes. *PLoS ONE*, **7**, e35039.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Altschul SF, Madden TL, Schäffer AA *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, **25**, 3389–3402.

Arnason U, Gullberg A, Widegren B (1993) Cetacean mitochondrial DNA control region: sequences of all extant baleen whales and two sperm whale species. *Molecular Biology and Evolution,* **10**, 960–970.

Arnason U and Gullberg A (1994). Relationship of baleen whales established by cytochrome b gene sequence comparison. *Nature* **367**, 726-8

Baker CS, Palumbi SR (1994) Which whales are hunted? A molecular genetic approach to monitoring whaling. *Science*, **265**, 1538–1539.

Baker CS, Perry A, Bannister JL *et al.* (1993) Abundant mitochondrial DNA variation and world-wide population structure in humpback whales. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 8239–8243.

Benson DA, Cavanaugh M, Clark K *et al.* (2013) GenBank. *Nucleic acids research*, **41**, D36–42.

Caputo V, Giovannotti M (2009) Haplotype characterization of a stranded *Balenoptera physalus* (Linnaeus, 1758) from Ancona (Adriatic sea, central Italy). *Hystrix, the Italian Journal of Mammalogy*, **20**, 83–85.

Clayton D (2000) *Islands of Truth: The Imperial Fashioning of Vancouver Island*. UBC Press, Vancouver.

Cooper A, Poinar HN (2000) Ancient DNA: do it right or not at all. *Science*, **289**, 1139–1139.

COSEWIC (2011) COSEWIC assessment and status report on the Humpback Whale *Megaptera novaeangliae* in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. x + 32 pp.

COSEWIC (2013) COSEWIC status appraisal summary on the Sei Whale *Balaenoptera borealis*, Pacific population, in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. xii pp.

Coté C (2010) *Spirits of our Whaling Ancestors: Revitalizing Makah & Nuu-chah-nulth Traditions.* UBC Press, Vancouver.

Cumbaa SL (1986) Archaeological evidence of the 16th century Basque right whale fishery in Labrador. *Report of the International Whaling Commission (Special Issue 10)*, 187–190.

Dalebout ML, Mead JG, Baker CS, Baker AN, Helden AL (2002) A new species of beaked whale Mesoplodon perrini sp. n.(Cetacea: Ziphiidae) discovered through phylogenetic analyses of mitochondrial DNA sequences. *Marine Mammal Science*, **18**, 577–608.

Foote AD, Hofreiter M, Morin PA (2012) Ancient DNA from marine mammals: Studying long-lived species over ecological and evolutionary timescales. *Annals of Anatomy - Anatomischer Anzeiger*, **194**, 112–120.
Frasier TR, Koroscil SM, White BN, Darling JD (2011) Assessment of population substructure in relation to summer feeding ground use in the eastern North Pacific gray whale. *Endangered Species Research*, **14**, 39–48.
Gravlund P, Aaris-Sørensen K, Hofreiter M *et al.* (2012) Ancient DNA extracted from Danish aurochs (Bos primigenius): Genetic diversity and preservation. *Annals of Anatomy*, **194**, 103–111.
Greenlee DM, Dunnell RC (2010) Identification of fragmentary bone from the Pacific. *Journal of Archaeological Science*, **37**, 957–970.
Grier C, Flanigan K, Winters M *et al.* (2013) Using ancient DNA identification and osteometric measures of archaeological Pacific salmon vertebrae for reconstructing salmon fisheries and site seasonality at Dionisio Point, British Columbia. *Journal of Archaeological Science*, **40**, 544–555.
Harkin M (1998) Whales, chiefs, and giants: an exploration into Nuu-chah-nulth political thought. *Ethnology*, 317–332.
Harris DJ (2003) Can you bank on GenBank? *Trends in Ecology and Evolution*, **18**, 317-319.
Huelsbeck DR (1988) Whaling in the pre-contact economy of the central northwest coast. *Arctic anthropology*, 1–15.
Huelsbeck DR, Fisken M (1983) *The Utilization of Whales at Ozette*. Laboratory of Anthropology, Washington State University, Pullman, Washington.
Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.
Indian and Northern Affairs Canada (2006) Maa-nulth First Nations Final Agreement (the "Final Agreement") Harvesting of Grey Whale and Sei Whale. Eric Denhoff, Chief Federal Negotiator.
Irwin DL, Mitchelson KR and Findlay I (2003) PCR product cleanup methods for capillary electrophoresis. *BioTechniques,* **34**,932-936.
Kaliszewska ZA, Seger J, Rowntree VJ, et al. (2005) Population histories of right whales (Cetacea: Eubalaena) inferred from mitochondrial sequence diversities and divergences of their whale lice (Amphipoda: Cyamus). *Molecular Ecology*, **14**, 3439–3456.
Keane TM, Creevey CJ, Pentony MM, Naughton TJ, Mclnerney JO (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evolutionary Biology*, **6**, 29.
Kool R (1982) Northwest Coast Indian whaling: new considerations. *Canadian Journal of Anthropology*, **3**, 31–44.
Koski LB and Golding GB (2001) The closest BLAST hit is often not the nearest neighbor. *Journal of Molecular Evolution,* **6**, 540-2.
Kubatko LS, Degnan JH (2007) Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, **56**,17–24.
Larkin MA, Blackshields G, Brown NP, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.
Malik S, Brown MW, Kraus SD, White BN (2000) Analysis of mitochondrial DNA

diversity within and between North and South Atlantic right whales. *Marine Mammal Science*, **16**, 545–558.
McLeod BA, Brown MW, Moore MJ *et al.* (2008) Bowhead whales, and not right whales, were the primary target of 16th-to 17th-century Basque whalers in the western North Atlantic. *Arctic*, 61–75.
McMillan AD, Claire DES (2005) *Ts'ishaa*. Archaeology Press. Department of Archaeology, Simon Fraser University, Burnaby.
McNeely R, Dyke AS, Southon JR (2006) Canadian marine reservoir ages, preliminary assessment. Geological Survey of Canada Open File 5049.
Monks GG (2003) The Cultural Taphonomy of Nuu-chah-nulth Whale Bone Assemblages. In: *Emerging from the Mist: Studies in Northwest Coast Culture History* (eds Matson RG, Coupland G, Mackie Q), pp. 188–212. UBC Press, Vancouver.
Monks GG, McMillan AD, St Claire DE (2001) Nuu-chah-nulth whaling: archaeological insights into antiquity, species preferences, and cultural importance. *Arctic anthropology*, 60–81.
Mulligan CJ (2006) Anthropological applications of ancient DNA: Problems and prospects. *American Antiquity*, 365–380.
Naidu A, Fitak RR, Munguia-Vega A, Culver M (2011) Novel primers for complete mitochondrial cytochrome b gene sequencing in mammals. *Molecular Ecology Resources*, **12**, 191–196.
Newman ME, Parboosingh JS, Bridge PJ, Ceri H (2002) Identification of Archaeological Animal Bone by PCR/DNA Analysis. *Journal of Archaeological Science*, **29**, 77–84.
Nichol LM, Gregr EJ, Flinn R *et al.* (2002) *British Columbia Commercial Whaling Catch Data 1908 to 1967 : A Detailed Description of the B.C. Historical Whaling Database*. Fisheries and Oceans Canada, Nanaimo.
Parson W, Pegoraro K, Niederstätter H, Föger M, Steinlechner M (2000) Species identification by means of the cytochrome b gene. *International journal of legal medicine*, **114**, 23–28.
Pastene LA, goto M, kanda N, et al. (2007) Radiation and speciation of pelagic organisms during periods of global warming: the case of the common minke whale, *Balaenoptera acutorostrata*. *Molecular Ecology,* **16**, 1481–1495.
Rastogi T, Brown MW, McLeod BA *et al.* (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**, 1647–1654.
Reeves RR (2002) The origins and character of "aboriginal subsistence"whaling: a global review. *Mammal Review*, **32**, 71–106.
Ronquist F, Teslenko M, van der Mark P *et al.* (2012) MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology*, **61**, 539–542.
Rooney AP, Honeycutt RL, Derr JN (2001) Historical population size change of bowhead whales inferred from DNA sequence polymorphism data. *Evolution*, **55**, 1678–1685.
Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings in the National Academy of Sciences of the United States of America*, **74**, 5463–5467.
Sasaki T, Nikaido M, Hamilton H, et al. (2005) Mitochondrial phylogenetics and evolution of mysticete whales. *Systematic Biology*, **54**, 77–90.

Schmidt HA, Strimmer K, Vingron M, Haeseler von A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*, **18**, 502–504.
Sinding M-HS, Gilbert MTP, Grønnow B *et al.* (2012) Minimally destructive DNA extraction from archaeological artefacts made from whale baleen. *Journal of Archaeological Science*, **39**, 3750–3753.
Taylor BL, Chivers SJ, Larese J, Perrin WF (2007) Generation length and percent mature estimates for IUCN assessments of cetaceans. *Southwest Fisheries Science Center*.
Thompson JD, Gibson TJ, Plewniak F, et al. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**, 4876–4882.
Torres-Florez JP, Hucke-Gaete R, Rosenbaum H, Figueroa CC (2014) High genetic diversity in a small population: the case of Chilean blue whales. *Ecology and Evolution*, **4**, 1398–1412.
Villesen P (2007) FaBox: an online toolbox for fasta sequences. *Molecular Ecology Notes*, **7**, 965–968.
Yang DY, Watt K (2005) Contamination controls when preparing archaeological remains for ancient DNA analysis. *Journal of Archaeological Science*, **32**, 331–336.
Yang DY, Woiderski JR, Driver JC (2005) DNA analysis of archaeological rabbit remains from the American Southwest. *Journal of Archaeological Science*, **32**, 567–578.
Yoshida H, Yoshioka M, Shirakihara M, Chow S (2001) Population structure of finless porpoises (*Neophocaena phocaenoides*) in coastal waters of Japan based on mitochondrial DNA sequences. *Journal of Mammalogy* **82**, 123–130.

Chapter 3

Genetic analyses of ancient and contemporary samples infer how historic and pre-historic events have shaped modern whale populations.

Abstract

It is becoming more important for conservation biologist to be aware of the phenomena known as "shifting baselines" when it comes to setting management goals for those species that have been negatively impacted by human behaviour. Whales represent a broad group of mammals for which discovering their "unaffected" state has been difficult due to a lack of accurate historic record keeping during an intense period of commercial whaling in the late 19th and early 20th centuries. In the decades since the moratorium on commercial whaling the populations of some species have increased, but the degree of recovery is often unclear with regard to whether or not they have reached their pre-whaling population sizes. Contemporary (post- commercial whaling) and ancient (pre-commercial whaling) samples were collected from two species off the west coast of Vancouver Island, and genetic analyses were conducted to better understand the effects of whaling on these populations. Although both contemporary populations had significantly lower levels of genetic diversity than their corresponding pre-exploitation counterparts, genetic signatures specific for recent bottlenecks were not detected, likely due to insufficient time since whaling for it to leave such genetic signatures. However, Bayesian skyline plots based on the ancient samples did reveal pre-exploitation population trends, but these same patterns were not found in the contemporary data, indicating that whaling erased the genetic signature of pre-exploitation population trends. In addition to improving our understanding of the history of these particular whale species, this work also has broad implications, showing that studies inferring population histories from contemporary samples need to be aware of the potential for recent dramatic changes to demography (e.g., from human exploitation) to erase previous signatures, leading to incorrect assessments of population history, and potentially contributing to the shifting baseline syndrome.

Introduction

Whales in every ocean around the world were negatively impacted by anthropological innovation and technology when it came to advances in whaling, with commercial processing ships being the most detrimental (Nichol *et al.* 2002). By the mid 1900s, whales were so depleted throughout the Earth's oceans that the International Whaling Commission was formed and a moratorium was placed on all commercial whaling in 1986, which is recognized by those countries that are members of the Commission (Reeves 2002; Coté 2010; Rose *et al.* 2011). The purpose of this moratorium is to give whales species a chance to recover, and to decrease the

likelihood of extinction due to over hunting (for those that are not already extinct). There is still legal whaling for sustenance throughout many places of the world, and Japan legally, takes some whales each year for "scientific" reasons, though this is surrounded by controversy. However, the number of whales taken annually is not nearly as many as were taken in the past. Since this stoppage of commercial whaling some populations have increased in numbers, such as the eastern North Pacific gray whale (*Eschrichtius robustus*) and the North Pacific humpback whale (*Megaptera novaeangliae*). These two populations have recovered to the point of being delisted from American and Canadian endangered species lists (COSEWIC 2011). Although there has been an increase in whale numbers for these species, there is still the problem of large-scale illegal whaling that happens around the world (e.g. Robert *et al.* 2009), as well as incidental mortalities from ship strikes and entanglements in fishing gear (e.g., Moore 2014), and because of these issues little is known about the true number of whales killed each year due to human actions. Additionally, it is still unknown if these whales have reached their pre-exploitation population sizes, or indeed what reasonable estimates of their pre-exploitation sizes are. This uncertainty surrounding population numbers and whaling's true impact, and the impacts on the context within which contemporary populations are interpreted, has led many researchers to find other ways to infer population history (Chan *et al.* 2005; Alter *et al.* 2007; McLeod *et al.* 2008; 2010; Foote *et al.* 2012; Alter *et al.* 2012).

Historical Demography and Genetics

The demographic history of a population can be inferred from the genetic

signatures found within the modern representatives (Alter *et al.* 2007; Ho & Shapiro 2011; Frasier *et al.* 2011; Palsbøll *et al.* 2012). In some cases, researchers may be able to get substantial quantity and quality of information from examining the genetic information from relatively few individuals representing the extant population. Genetic signatures are patterns or changes found in either the mtDNA or nDNA that can provide information on what events have occurred that may not have been observed through physical monitoring. The examination of such signatures have been used to better understand the history of a wide range of species, such as the endangered red pandas (Hu *et al.* 2011) and blue whales (Torres-Florez *et al.* 2014).

Additionally, the genetic signatures found within ancient samples can also be helpful in exploring historic demography of species before human impact (Nyström *et al.* 2012). The DNA from the now extinct cave bears (Stiller *et al.* 2010) and woolly mammoths (Barnes *et al.* 2007; Nyström *et al.* 2012) was used to inferred what effect humans and climate change had on these species. These signatures in ancient samples can also be compared to contemporary samples to investigate changes over time (Rooney *et al.* 2001; Lorenzen *et al.* 2011; Alter *et al.* 2012). This approach is becoming more and more important in conservation biology to prevent the incidences of "shifting baselines" (Pauly 1995; Sheppard 1995). "Shifting baselines" can occur when the unaltered state of a population is unknown, and current management decisions are based upon data that may be biased towards that which is available (Pauly 1995). This means that what may be considered the "normal" state for a population is actually an altered state, for example, the population has decreased in the past, but this has gone undetected and therefore it is though the population size was always small. Historical events that are unidentified or

occurred before the beginning of data collection can have a large influence on management, and genetic signatures from ancient samples can assist in ensuring the correct decisions are made, as ancient samples may provide evidence of this "normal" or unaltered state of a population

One type of historical event that can leave a strong signature in the genetic characteristics of a population is that of a population decline or bottleneck event (Peery *et al.* 2012). Based on theory, and the examination of scenarios of known population declines and severe bottlenecks, conservation biologist and population geneticist can develop a reference on how the genetic variation may change as a result of this loss in individuals and their genetic contribution, and what influence it may have on the species or population as a whole (Tajima 1989; Fu & Li 1993; Luikart & Cornuet 1998; Le Page *et al.* 2001; Garza & Williamson 2001; Bellinger *et al.* 2003). They can then use this information to infer historical demography, and to predict future responses to changes in habitat or climate, for those species that have not been as well documented in the past (Lorenzen *et al.* 2011). This reference can also be helpful for populations in which long-term monitoring was not possible, physically or financially, because examining this genetic variation may provide as much information as long term monitoring (Balazs & Chaloupka 2004; Stiller *et al.* 2010; Nyström *et al.* 2012).

Historic Demography of Whales

Traditionally, the historic demography of populations prior to and during the time of commercial whaling was inferred through the use of whaling ship logbooks, and records regarding the export of whale byproducts, such as barrels of oil and baleen (Smith &

Reeves 2003; Carroll *et al.* 2014). These methods are helpful, but incomplete records can mean an over or under estimation of how many individuals where once present, and it can also lead to issues if the incorrect species where recorded. If a logbook recorded the wrong species in their whaling period then this could mean that what was once thought of as an abundant species, because so many where being taken, was actually already a small population. Investigating the genetic signatures found within the nuclear and mitochondrial DNA can be used assist these previously utilized methods in inferring the historic and pre-historic demography of those whales affected by the period of intense commercial whaling, such as the gray and humpback whales found within the North Pacific Ocean.

Objectives

The objectives of this project were to examine the DNA from contemporary and ancient representatives of gray and humpback whales that summer off the west coast of Vancouver Island to test for such genetic signatures and therefore obtain a better understanding of pre-exploitation trends, as well as the impact of whaling. The contemporary samples from these species were used to examine for the signatures associated with a recent bottleneck event, which was the result of commercial whaling. The ancient samples, or pre-commercial whaling samples, were used to explore what affect this large loss of individuals had on the population by: (a) comparing to contemporary samples to reveal if there was significant loss in genetic variability due to exploitation; and (b) to infer population demography prior to commercial whaling and investigate any possible differences in the estimations past population size, as this

could affect current and future recovery goals. The use of both contemporary and ancient samples will help researchers to see a more complete picture of the demographic history of gray and humpbacks that summer of the west coast of Vancouver Island, and hopefully prevent the phenomena of "shifting baselines".

Materials and methods

Contemporary sample collection and DNA extraction

Many whale species show site fidelity to feeding or calving grounds, meaning that they will return to the same site over many years or even throughout their lifetime, and will pass this preference for an area to their offspring (Cypriano-Souza *et al.* 2010; Frasier *et al.* 2011). For this reason, contemporary and historic samples were collected from similar locations to ensure that data from the two time periods were comparable, therefore all contemporary skin samples were collected via modified crossbow from identified gray and humpback whales that frequent the west coast of Tofino, British Columbia during the summer feeding months (e.g. Lambertsen 1987; Palsbøll et al. 1991; Frasier *et al.* 2011). These tissue samples were stored in a 20% dimethyl sulfoxide (DMSO) solution and transported back to Saint Mary's University in Halifax for DNA extraction (Seutin et al. 1991).

The DNA extraction process, from ~40mg of contemporary skin tissue, was completed following a standard phenol:cholorform method that is commonly used for tough tissues, such as whale skin (Wang *et al.* 2008). The quantity of DNA obtained from each sample extracted was estimated based on spectrophotometry using a NanoDrop 2000 (Thermo Scientific Inc.). The quality of the extracted DNA, to determine if there was

any DNA degradation, was assessed based on electrophoresis of 25 ng of DNA through 2.0% agarose gels stained with SYBR Green I (Invitrogen).

Ancient sample collection and DNA extraction

Previously identified ancient gray and humpback whale samples were also collected on and around the coast of Tofino, BC. These samples were from whale bones that were recovered from documented First Nations traditional whaling sites (Clayton 2000) and were later identified to species through genetic methods (see Chapter 2). Bone shavings and smaller bone fragments were collected and then transported back to Saint Mary's University for DNA extraction. As part of a related ongoing project, radiocarbon dating was conducted on 5 of the bones. The purpose of dating these samples was to provide a rough time frame for bones in the area, and to demonstrate that those found on the surface of the sampling sites could in fact be classified as ancient. The 5 bones that were selected were those samples with enough shavings (1-2 grams) available without compromising the amount necessary for species identification.

DNA was extracted from 0.15-0.18 grams of bone shavings using a modified Qiagen column extraction as per Rastogi *et al.* (2004) and McLeod *et al*. (2008). Whaling traditions on the West coast of Vancouver Island date back thousands of years, therefore the DNA from recovered bones was likely of low quality and quantity, and highly susceptible to contamination from contemporary samples (Mulligan 2006). Thus, special protocols needed to be followed when working with ancient DNA (aDNA). For example, as suggested by Cooper and Poinar (2000), all aDNA processing, prior to PCR amplification, was carried out in a dedicated and isolated laboratory where analysis on

extant cetaceans has never been preformed, and negative controls were included at each step of extraction and amplification to ensure there was no contamination (Yang & Watt 2005).

mtDNA Sequencing and Analyses

For both ancient and contemporary samples, a portion of the mitochondrial DNA control region was amplified and sequenced. A ~450 bp region of the mitochondrial control region was amplified using the primers t-PRO and Primer-2 from Yoshida *et al.* (2001). This was conducted on 47 contemporary gray whale and 22 humpback whale tissue samples. The DNA extract was diluted down to 5ng/μL with $TE_{0.1}$. PCR amplifications were conducted in 20μl volumes containing: 2μL of 5ng/μL DNA dilution, 1X PCR buffer (Promega Inc.), 1.5mM $MgCl_2$, 0.2mM each dNTP (Invitrogen), 0.16μg/μL bovine serum albumin (BSA), 0.05U/μL *Taq* polymerase (Promega Inc.), 0.3μM of each primer. PCR cycling conditions for the control region consisted of an initial 5 minute denaturation step at 94°C; 30 cycles of 94°C for 30 seconds, 55°C for one minute, and 72°C for one minute; with a final extension step at 60°C for 45 minutes. Thirty-two previously identified ancient gray and 27 ancient humpback whale bones samples were amplified and sequenced using the same primers. PCR amplifications were also conducted in 20μl volumes containing: 5μL DNA extract, 1X PCR buffer (ProMega Inc.), 1.5mM $MgCl_2$, 0.2mM each dNTP (Invetrogen), 0.3μg/μL bovine serum albumin (BSA), 0.1U/μL *Taq* polymerase (Promega Inc.), and 0.3μM of each primer. The PCR cycling conditions for the amplification of the control region were different, as they were conducted using touchdown PCR. Touchdown PCR consisted of an initial 5 minute

denaturation step at 94°C; 50 cycles of 94°C for 30 seconds, 65-55°C for one minute (see below), and 72°C for one minute; with a final extension step at 60°C for 45 minutes. The starting annealing temperature was 65°C, which was lowered by 0.5°C every cycle until reaching 55°C, where it remained for 30 cycles. All PCR products, both contemporary and ancient, were then visualized via electrophoresis through 1.5% agarose gels stained with ethidium bromide.

In preparation for sequencing PCR products were purified to remove excess dNTPs and primers that were not incorporated during the initial amplification, as they could interfere with the sequencing reaction (Werle *et al*. 1994). This process was conducted in an enzymatic reaction using reagents from New England BioLabs®, which required 0.65ul Antarctic Phosphatase Buffer, 0.1µl Antarctic Phosphatase, and 0.03µl Exonuclease I per 5µl sample of PCR product. The thermal cycling conditions were 15 minutes at 37°C, followed by 15 minutes at 80°C and finally held at 10°C.

The ancient DNA samples were sequences in both the forward and reverse directions for both regions to ensure that correct base pair assignment and to create consensus sequences, whereas the contemporary samples were only sequenced in one direction. The sequencing PCR cocktail, with all reagents from Applied Biosystems®, took place within 15µl reactions and contained 0.25X Big Dye Terminator reaction mix, 1X Big Dye Terminator sequencing buffer, 1µl/reaction of 5µM of specified primer, and 5µl/reaction of 4.5ng/µl control region PCR product. The PCR cycles consisted of an initial 2 minute denaturation step at 96°C; 30 cycles of 96°C for 20 seconds, 50°C for 20 seconds, and 60°C for 4 minutes; with a final hold step at 4°C.

Once the sequencing reaction was completed, an ethanol precipitation procedure

was conducted to de-salt the samples (Irwin *et al*. 2003). The resulting DNA was re-suspend in 10µl HiDi formamide and analyzed on an ABI 3500xL genetic analyzer for sequencing.

Before analyses could be conducted on the sequence data, 4Peaks software Version 1.7.2. (Griekspoor and Groothuis 2006) was used to edit sequences by eye to ensure correct base assignment. For the ancient samples, the forward and reverse compliment sequences for each sample were aligned in ClustalX (Thompson *et al.* 1997; Larkin *et al*. 2007) and a consensus sequences was created. Once all the sequences were confirmed, the various analyses could be preformed on the ancient and contemporary samples independently or together, depending on the analysis.

Control region haplotypes and their frequencies, within each species, were determined using Fabox v1.4 (Villesen 2007). One hundred and sixty-one contemporary gray whale control region sequences that were previously obtained by Frasier *et al.* (2011) were added to the gray whale network. Twenty-eight contemporary control region sequences from Baker *et al.* (2013) were added to the humpback whale network. The addition of these previously published sequences was to ensure a wide representation on the haplotypes that are currently found in todays Pacific whale populations. The script TempNet (Prost & Anderson 2011), which utilizes R functions (R Core Team 2014) to visualize haplotype relationships from FaBox v1.4, was used to create a parsimony network of the control region sequence data from two time periods for each species. As this network represents two time periods, they were designated as ancient (pre-commercial whaling) and as contemporary (post-commercial whaling). This was done to illustrate the possible changes in control region haplotypes over time and due to

commercial whaling.

The control region sequences of the contemporary gray and humpback whales were evaluated according to a mismatch distribution to look for genetic signatures of expansion, which can follow a bottleneck event (Rogers & Harpending 1992). A recent population expansions will result in a smooth unimodal distribution due to the star-like phylogeny of the sequences being compared, meaning there will be an increase in the frequency of sequences with lower amounts of pair-wise differences and fewer with larger amounts of pair-wise differences (Rogers & Harpending 1992). The raggedness index, which indicates the fit of the data to the model of population expansion, was used to check the significance of the mismatch distributions and should be low for those distributions that demonstrate expansions (Harpending et al. 1993; Harpending 1994). The program ARLEQUIN V3.5.1.2 (Excoffier *et al.* 2005; Excoffier & Lischer 2010) was used to calculate the mismatch distributions and the tests for additional signatures of a bottleneck event or population expansion, such as Tajima's *D* and Fu's *Fs* (Tajima 1989; Fu 1997). Tajima's *D* is often used to infer long-term changes in effective population size by comparing the number of segregating sites with nucleotide diversity in mtDNA sequences (Tajima 1989; Peery *et al.* 2012). Those populations that have undergone a bottleneck are expected to have a positive Tajima's *D* because the number of segregating sites will be reduced more quickly during a bottleneck than the average per nucleotide diversity. Fu's *Fs* will infer if there has been a population growth, such as one that would happen after a bottleneck event (Fu 1997; Bos *et al.* 2008). A negative value of Fu's *Fs* is evidence an excess of rare alleles, as would be expect from a recent population expansion because the test is based on the probability of observing more alleles in a sample of a

given size, conditioned on the observed average number of pairwise differences (Fu 1997; Excoffier *et al.* 2005; Excoffier & Lischer 2010). These analyses were done to investigate if enough time had passed since the end of commercial whaling to see any genetic signatures in the mtDNA. Only those contemporary samples obtained from the west coast of Vancouver Island, either during the course of this project or previously reported in Frasier *et al.* (2011) were used for these analyses, to ensure they were all from the same sampling area.

nDNA Microsatellites and Analyses

The contemporary samples were amplified at 15 target microsatellites in 5 separate multiplex reactions (Table 3.1) (D'Intino *et al.* 2013) . All primer pairs were ordered with one of four fluorescent tags (6FAM- Blue, VIC-Green, PET-Red, or NED-Yellow). PCR amplification for the 5 multiplex reactions were carried out in 20 µL reaction volumes containing 10 ng template DNA, 1X PCR buffer (Promega Inc.), 0.2 mM of each dNTP (Invitrogen), 1.5 mM $MgCl_2$ (Promega Inc.), 0.15 - 0.75 µM of each primer (D'Intino *et al.* 2013), 0.16 mg/mL of bovine serum albumin (BSA), and 0.05 U/µL of *Taq* polymerase (Promega Inc.). PCR cycling conditions consisted of an initial denaturing step of 95^{o}C for 5 min; 30 cycles of denaturation at 95^{o}C for 30 sec, annealing at 55^{o}C for 1 min, 72^{o}C for 1 min; and a final extension period at 64^{o}C for 45 min. PCR product was combined with HiDi^{TM} formamide and GeneScan-600 LIZ size standard (Life Technologies), and finally size-separated. Samples were then separated by size and visualized on an Applied Biosystems 3500xL Genetic Analyzer.

For the aDNA samples, the five microsatellite loci: RW31, FCB17, TexVet5,

EV37Mn and GATA417, were selected from the 15 microsatellites used for in the contemporary samples, and were chosen based on their ability to amplify well in both gray and humpback whale samples (Table 3.1) (D'Intino *et al.* 2013). Each locus was amplified separately in 5 independent reactions. PCR cycling conditions were the same as those described for aDNA mitochondrial control region amplification.

Similar to the work conducted by McLeod *et al.* (2010) to account for potential genotyping problems in aDNA due to low template and allelic dropout (Taberlet et al. 1996), up to eight replicate amplifications were conducted on the ancient samples to ensure that both alleles were detected. A locus was determined to be heterozygous if each allele was observed at least twice and homozygous if the profile at a particular locus was independently determined as homozygous seven times (Taberlet et al. 1996; McLeod *et al.* 2010). Unlike the contemporary samples, which were amplified as part of a multiplex reaction, the ancient samples were amplified separately at each locus and then visualized via electrophoresis through 1.5% agarose gels stained with ethidium bromide (EtBr) independently for all 5 loci. This was to confirm amplification of the aDNA samples. PCR products were first diluted based on the EtBr agarose gel to ensure that the appropriate amount was used in subsequent analysis, as too little or too much can cause scoring errors, then the PCR product was combined with HiDi^{TM} formamide and GeneScan-600 LIZ size standard (Life Technologies), and finally size-separated. Like the contemporary samples, they were then separated by size and visualized on an Applied Biosystems 3500xL Genetic Analyzer.

Each locus, both within the 5 multiplex reactions and the independent loci, were examined by eye and allele sizes were calculated using GeneMarker software 2.0

(SoftGenetics).

To examine for the genetic evidence of a recent population bottleneck in the microsatellites of the contemporary gray and humpback sample sets, which is well documented due to the period of intense commercial whaling of both species, the analytical method explained in Cornuet and Luikart (1996) and employed in the program BOTTLENECK v. 1.2.02 (Piry *et al.* 1999) was used. This method tests for 'heterozygosity excess' based on a comparison between observed and expected heterozygosity (Peery *et al.* 2012). This excess occurs because when a population has a reduction in population size, there is a reduction in the number of alleles and heterozygosity a polymorphic loci, and the allele numbers is reduced faster than the heterozygosity (Nei *et al.* 1975; Piry *et al.* 1999). Additionally, BOTTLENECK v. 1.2.02 examines the mode-shift of the alleles, where a bottleneck is detected when one or more of the common allele classes have a higher number of alleles than the rare allele class, therefore deviating from the normal L-shape (Luikart *et al.* 1998). This is because a bottleneck event will cause alleles at low frequency to become less abundant than alleles in one or more intermediate allele frequency class (Luikart *et al.* 1998; Koringa *et al.* 2008). As suggested by Piry et al. (1999), a two-phase mutation (TPM) model was used with the following parameters: 95% single-step mutations, 5% multistep mutations, and 10000 iterations. A one-tailed Wilcoxon signed-rank test was performed to test for significance of heterozygosity excess due to only having 15 loci (Wilcoxon 1945).

Additionally, the microsatellites were used to search for genetic evidence of a population bottleneck in the two species with the *M*-ratio analysis, which is based on the ratio *k/r*, in which *k* is the total number of alleles and *r* is the overall range in allele

size (Garza & Williamson 2001). Therefore, the ratio is expected to be smaller in those species/populations that have experienced recent bottlenecks than those that have remained constant because the number of alleles will be reduced faster than the allele range (Garza & Williamson 2001; Peery *et al.* 2012). The rarest alleles are usually lost first in a bottleneck event but reduction of size range in alleles (*r)* would only be reduced as quickly as the number of alleles (k) if the rarest alleles were the largest and smallest, which isn't often the case (Garza & Williamson 2001). The mean observed *M*-ratio (across multiple loci) was compared to the calculated critical value, which was based on a population in equilibrium being simulated 10 000 times and *M*-ratio calculated for each replicate. The values of *M* were sorted in descending order and the critical value (*Mc*) was determined such that 95% of the simulations of an equilibrium population had $M > Mc$ (Garza & Williamson 2001). If $M < Mc$ than a population bottleneck can be inferred. The two programs, M_P_VAL and CRITICAL_M, were used to examine this ratio (Garza & Williamson 2001). The analysis was run varying the following parameters: the range of estimated pre-commercial whaling population sizes for each species (Baker *et al.* 1993; Alter *et al.* 2012), average size of multi-step mutations (Δg) and proportion of one-step mutations (*ps*) as recommended by Garza and Williamson (2001) and Peery *et al.* (2012) (Table 3.2) (Torres-Florez *et al.* 2014). The analysis was run at varying parameters to attempt to encompass the true values of Δg and *ps* for gray and humpback whales. A typical mutation rate of $\mu = 5x10^{-4}$ for mammalian microsatellites, which has been used in similar bottleneck testing for blue whales, was used (Torres-Florez *et al.* 2014).

Comparisons of allelic diversity and heterozygosity between ancient and

contemporary sample sets were conducted in R (R Core Team 2014). Due to the differences in samples sizes between the ancient and contemporary data sets, a resampling method was used for standardization. Briefly, for each comparison sample sizes of 15 individuals were randomly selected from each of the contemporary and ancient sample sets. This randomization and resampling was repeated 100 times to generate comparable distributions of allelic diversity and heterozygosity between the two time periods (for each species). For this method allelic diversity was defined as the number of unique alleles at a locus divided by the number of alleles scored, in order to standardized for missing data, particularly in the ancient data sets; and heterozygosity analyses were based on observed heterozygosity. To test for significant differences in both the allelic diversity and heterozygosity results, the Wilcoxon signed-rank test was used, and the effect size was reported (Wilcoxon 1945; Kelley & Preacher 2012).

Inferring past demography

Past population size changes for gray and humpback whales were inferred based on the control region sequences and using a Bayesian coalescent approach employed in BEAST (Drummond & Rambaut 2007; Drummond *et al.* 2012). This program is able to estimate the genealogy of a population from the available sequence data and then estimate the historic demography based on that genealogy (Drummond 2005; Ho & Shapiro 2011; Drummond *et al.* 2012). It can estimate the overall effective population size through time from the sequences given, by providing an effective population size estimate at each coalescent point until the most recent common ancestor is reached (Chen *et al.* 2013). The ancient and contemporary samples where run independently to examine for differences

in the skyline plots, and therefore examine for difference in the estimation of demography. To prepare the sequences for this analysis, they were aligned with ClustalX and checked by eye using a range of gap opening and extension penalties, to ensure that alignments correctly accounted for such changes as additions or deletions in the sequences. The program eBioX was used to cut the sequences to the same number of base pairs (Martínez Barrio *et al.* 2009). The program ModelGenerator v85 (Keane *et al*. 2006) was used to identify the most appropriate model of molecular evolution and then TREE-PUZZLE (Schmidt *et al*. 2002) was used to estimate the transition:transvertion ratio, as well as the α value describing the shape of the gamma distribution if there was found to be rate heterogeneity in substitution rates across sites.

There are multiple priors for the reconstruction of the genealogies and population size change estimates, such as mutation model and mutation rate (Table 3.3), but the key variable that was tested was the mutation rate of the control region in cetaceans. Two mutation rates utilized and compared were 4.0×10^{-8} substitutions per site per year (Alter *et al.* 2012) and 3.5×10^{-9} substitutions per site per year (Baker *et al.* 1993). Alter *et al.* (2012) obtained their estimation of the control region mutation rate based on the cytochrome-*b* gene of gray whales, whereas Baker *et al*. (1993) calculated theirs control region mutation rate through divergent rate estimate of cetaceans. These two mutation rates were tested as they differ by almost an order of magnitude, which will not affect the overall shape or pattern of the genealogy but will influence the timing of the changes in demography. By using both mutation rates, it could be inferred that the true demography is found between. Additionally, by using the Alter *et al.* (2012) mutation rate, the results would be comparable to previous research. The Bayesian skyline plots were

reconstructed within TRACER v1.6 (Rambaut & Drummond 2007).

Results

Classification of Ancient Samples

Age estimates of the 5 dated samples ranged from 499-1139 years before present (BP=1950). This dating has been corrected for the Marine Reservoir Effect for the area (McNeely *et al*. 2006). The age range of the bones pre-date commercial whaling, and therefore are classified as ancient for the purpose of this study.

Mitochondrial DNA Analyses

The control region sequences of 47 gray and 22 humpback whales were successfully recovered from the contemporary tissue samples and 32 gray and 27 humpback whale control region sequences were obtained from the ancient bone samples. The contemporary and ancient sequences, for each species, were combined to create the two haplotype networks in order to observe the change in haplotype availability and frequency between the two designated time periods (Figure 3.1 and 3.2).

The mismatch distributions for the control region of the contemporary sample sets, for both the gray and humpback whale, had smooth and unimodal distributions which are associated with population expansion (Figure 3.3). The raggedness indices for the gray and humpback whale mismatch distributions were 0.011 (p-vaule = 0.52) and 0.0494 (p-value = 0.859), respectively. These low raggedness indices and p-vaules > 0.05 indicate that the data is a "good fit" to the model of expansion and the null hypothesis of expansion cannot be rejected (Harpending et al. 1993; Harpending

1994). Tajima's *D* was 0.77924 (p-value > 0.05) and -1.80192 (p-value < 0.05) for gray and humpback whales respectively. A positive value would indicate a population bottleneck, whereas a negative value would be indicative of a population expansion. The Tajima's *D* for the contemporary humpback population was significant for a population expansion. Fu's *Fs* was -3.7376 (p-value > 0.05) and 0.95246 (p-value > 0.05) for gray and humpback whales respectively. A positive value would indicate bottleneck and negative would indicate expansion, but neither of these were significant.

Microsatellite Analyses

The genotyping of the contemporary DNA from gray whales was successful for the 15 loci (189 individuals), but only for 14 loci with the contemporary humpback whales (22 individuals) (Table 3.1). The ancient DNA samples were only successfully genotyped at 4 of the original 5 loci. Locus RW31 was not successful when amplified in the ancient gray and humpback samples. When these microsatellite data were analyzed for genetic signatures associated to a bottleneck event, the Wilcoxon sign test for heterozygosity excess under the TPM did not lend support for either the gray (p-value = 0.35986) or humpback (p-value = 0.55371) whales having gone through a recent bottleneck event. Additionally, the allele frequency distribution (mode-shift) was L-shaped for both species, which is expected for non-bottleneck populations. When the *M*-ratio analysis was performed to also examine for signs of a population bottleneck, it provided little support. Under all scenarios for each species (i.e. range in Ne, Δg, and *ps*) the observed M-ratios were larger than the calculated Mc (Table 3.3). Additionally, all *M*-ratios were over the suggested threshold of 0.68 identified by Garza and Williamson

(2001) for bottlenecked populations.

The allelic diversity and heterozygosity of both species were significantly higher in the ancient samples than in the contemporary samples. Results for allelic diversity are shown in Figure 3.4a and 3.4b), with the Wilcoxon signed-ranked test resulting in p-values of 1.506×10^{-15} and 2.2×10^{-16}, and effect sizes of 1.30 and 2.22, for the gray and humpback whales respectively. When heterozygosity of the ancient and contemporary samples was compared, again both gray and humpback whales demonstrated significantly higher heteroygosity in the ancient samples (Figure 3.5a and 3.5b). The Wilcoxon signed-ranked test gave p-values of 2.2×10^{-16} and 1.098×10^{-11}, and effect sizes of 2.30 and 1.09, for the gray and humpback whales, respectively.

Inferring past demography

The inferences of the demographic history were visualized through the Bayesian skyline plots. Two skyline plots were constructed for each of the data sets: 1) ancient gray whale; 2) contemporary gray whale; 3) ancient humpback; and 4) contemporary humpback. When the gray whale ancient and contemporary sequence data were compared it resulted in two different patterns of historic demography. The skyline plot for the ancient gray whale data demonstrated a pattern of increasing population size beginning ~75,000 years in the past before stabilizing ~25,000 years ago (Figure 3.6a), whereas, the skyline plot utilizing only the contemporary gray whale sequence data showed no such increase, but only what appeared to be a stable population size throughout the past 400,000 years (Figure 3.6b). The overall demographic pattern observed was similar when the ancient and contemporary humpback whale data were

compared. Both inferred a stable population size, but the logarithm of the effective female population size and generation time was larger when the ancient data was used (Figure 3.6c-d). When the two different mutation rates used in the estimations were compared, they demonstrated the same pattern of change in demography, but the overall logarithm of the effective female population size and generation time was lower and in the skyline plots utilizing Atler *et al.* (2012) mutation rate for the control region (Figure 3.6a-d). This means that the estimation of historic and pre-historic populations size would be lower when using the Atler *et al.* (2012) mutation rate.

Discussion

Though both gray and humpback whale populations found in the North Pacific Ocean were drastically reduced due to the period of intense commercial whaling in the late 19th and early 20th centuries, the consensus from examining the contemporary nDNA is that there is a lack of genetic signatures indicating the known recent bottleneck event. The analyses of the mtDNA do suggest a population expansion, such as one that can be seen after a bottleneck event, demonstrating that this molecular marker may be more sensitive than bottleneck detection. The mtDNA indicated a population expansion at some point in the past, and examining the timing of these expansions and why they may differ between the gray and humpback whale populations will be examined in the future of this project. Both of the mismatch distributions did have a unimodal distribution, which is usually seen in those populations that have experienced a population expansion in the past and the raggedness indices indicated that the data was a good fit of the model of population expansion (Rogers & Harpending 1992; Johnson *et*

al. 2007; Torres-Florez *et al.* 2014). All but one of the tests for neutrality and signs of expansion on the mtDNA were insignificant (p-value > 0.05). Though the significant negative Tajima's *D* result for the contemporary humpback whale is inductive of a population expansion, which may follow a less recent bottleneck, the sample size of 22 individuals must be taken into account when reporting confidence to this result (Lessios *et al.* 2001; Johnson *et al.* 2007; Sastre *et al.* 2010). Furthermore, the other tests for bottleneck detection were all negative or not significant for this sample set. The power to detect bottleneck signatures in contemporary samples is dependent and limited by many factors, such as i) time since event, ii) generation time of species, iii) sample size and iv) genetic marker used (Busch *et al.* 2007; Peery *et al.* 2012). This is because these analytical methods, *M*-ratio and heterozygosity excess, do have their own limitations. These methods are most powerful when they are used recently after a population experiences the bottleneck event, but not too recently, as *M*-ratio has low detection success 1-5 generations post bottleneck (Busch *et al.* 2007; Peery *et al.* 2012). This low detection success of *M*-ratio could explain the results seen in cetaceans studies, as they are long-lived and have longer generations times, and simply there may have not been enough passage in time for the DNA to reflect these changes in population size, even though microsatellites have a fast mutation rate than mtDNA (Taylor *et al.* 2007; Alter *et al.* 2007; 2012; Torres-Florez *et al.* 2014).

When examine the genetic data from two time points, in this case it was pre- and post- commercial whaling, the effects of commercial whaling become significant (Ramakrishnan *et al.* 2005; Peery *et al.* 2012). The allelic diversity and heterozygosity comparison demonstrated both were higher before commercial whaling for gray and

humpback whales in the North Pacific. These networks, which showed the distribution of the haplotypes temporally, are helpful to visualize how the genetic variation has changed. It can also demonstrate if there was a loss of rare haplotype as a result of the decrease in population size by observing a lack of these rare haplotypes in the contemporary population, such as haplotype 25, 27 and 29 in the gray whale sample set (Figure 3.1). These networks are also helpful to understand how limited the ancient sample sets are, and by increasing the number of ancient samples in future research it could be revealed which haplotypes are new mutations, and which are simply missing from the data set.

The Bayesian skyline plots from both the ancient and contemporary data sets demonstrate the most interesting results from this study (Figure 3.6a-d). The skyline plots of the contemporary data sets are able to infer a decline in population size close to time zero, which is present day (Figure 3.6b and 3.6d). This could indicate the bottleneck event that was caused by commercial whaling. These results also clearly demonstrated the importance of conducting analyses on sample sets from more than one time point, as large demographic events, such as a bottleneck, could change the overall inference of the historic demographic pattern (Peery *et al.* 2012). If management decisions regarding the conservation of the species were to be made from the contemporary data alone, it would appear as though the population sizes of both gray and humpback whales have stayed relatively constant over time leading up to a recent sharp decline. But when the skyline plots from the ancient sample sets are examined, particularly the gray whale plot, there is a very different demographic inference (Figure 3.6a). The ancient gray whale plot suggests that the population began increasing in size around 200,000 years before time zero (499-1139 years BP, based on radiocarbon dating) and then began to stabilize

before the on-set of large scale commercial whaling. It is possible that the large decline in population due to commercial whaling actually erased this signature of historical increase in gray whales. It has been proposed with other study species that bottleneck events can have this masking result (Lessios *et al.* 2001; Busch *et al.* 2007; Johnson *et al.* 2007; Smith *et al.* 2011). Furthermore, the estimated logarithm of the effective female population size and generation time were lower when only the contemporary data was utilized to make demographic inferences (Figure 3.6c-d). The skyline plots from the contemporary data inferred a lower estimated population size due to the significant decrease in genetic variation found within the contemporary populations compared to the ancient sample set (Figure 3.4 and 3.5). These findings only add to the concern of "shifting baselines" within present population management (Pauly 1995; Sheppard 1995). If ancient data is ignored or not factored into demographic inferences, then assessments will be based solely on current data and this could lead to inappropriate decisions.

The overall message that can be concluded from this investigation in bottleneck detection and inferring of historical demography is that no one method should be used alone (Busch *et al.* 2007; Sastre *et al.* 2010). Each method has its strengths and weaknesses, but when they are applied together the conclusion will have more overall confidence and more informed estimations could be made. Also, if data from multiple time periods are available, they can assist to ensure that genetic signatures are not lost due to various demographic events. The use of ancient samples within this study demonstrate how more than one time point can lead to new inferences on past demographic events. These ancient samples were key in uncovering a significant loss of allelic diversity and heterozygosity that may have gone unnoticed had contemporary samples been the only

time point used. Furthermore, when it comes to inferring historic demography, this study clearly shows the importance of utilizing ancient data, and that the conclusions made through the use of only contemporary data may not be ideal for all species, though this is the type of data that was and still is commonly used to make inferences on the past.

Humans have made a huge impact on the world, and when it comes to other species that we share it with, it is more often than not, a negative impact. When we look at long lives species, such as cetaceans, we are still discovering what effects commercial whaling had on their genetic composition and that we may not see the true results until many years to come (Waldick *et al.* 2002). Collecting as much genetic information as possible, from multiple time periods, may be the only way in which we are able to discover what effect past event have had on these whale species, and this will assist in making predictions on their future.

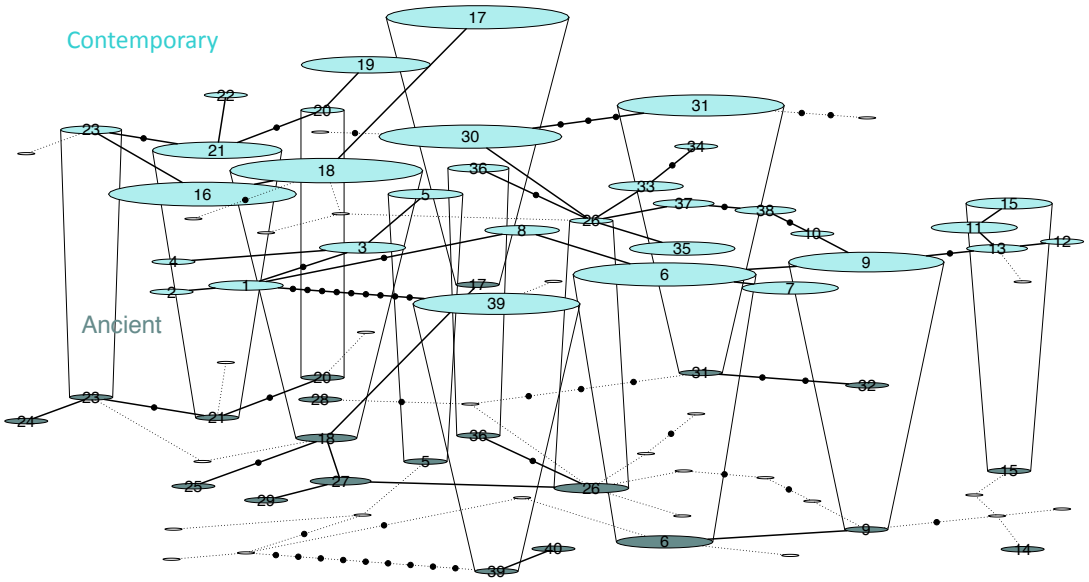

Figure 3.1. Parsimony network of control region sequence data for gray whale samples. The network for the contemporary samples is situated about the network for the ancient samples. Numbers within the ellipses indicate the haplotype designation and connection between time periods. Open circles indicate the haplotype is missing from one of the time periods, black circles indicate a mutation between the haplotypes.

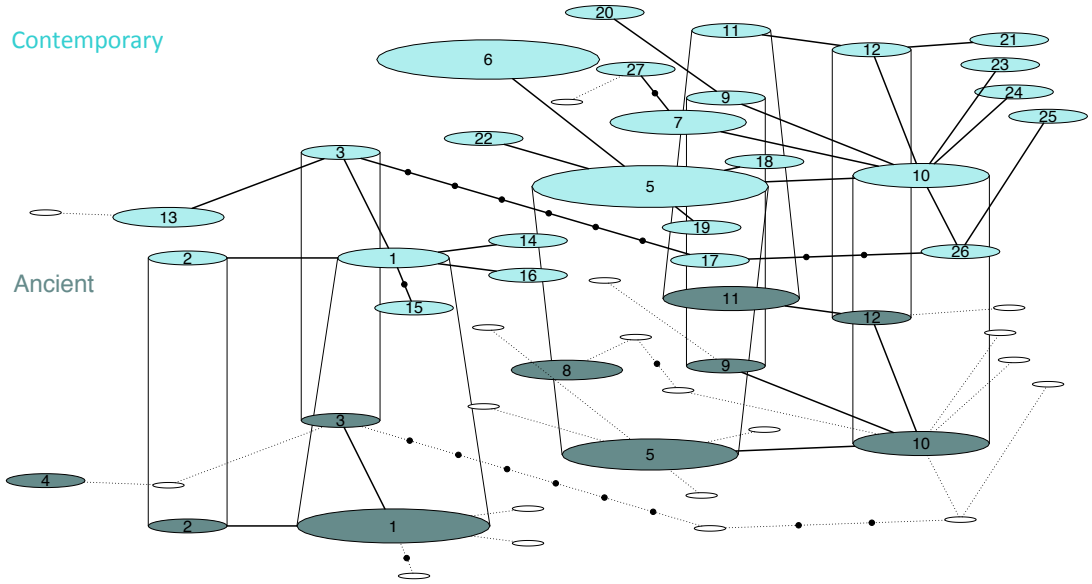

Figure 3.2. Parsimony network of control region sequence data for humpback whale samples. The network for the contemporary samples is situated about the network for the ancient samples. Numbers within the ellipses indicate the haplotype designation and connection between time periods. Open circles indicate the haplotype is missing from one of the time periods, black circles indicate a mutation between the haplotypes

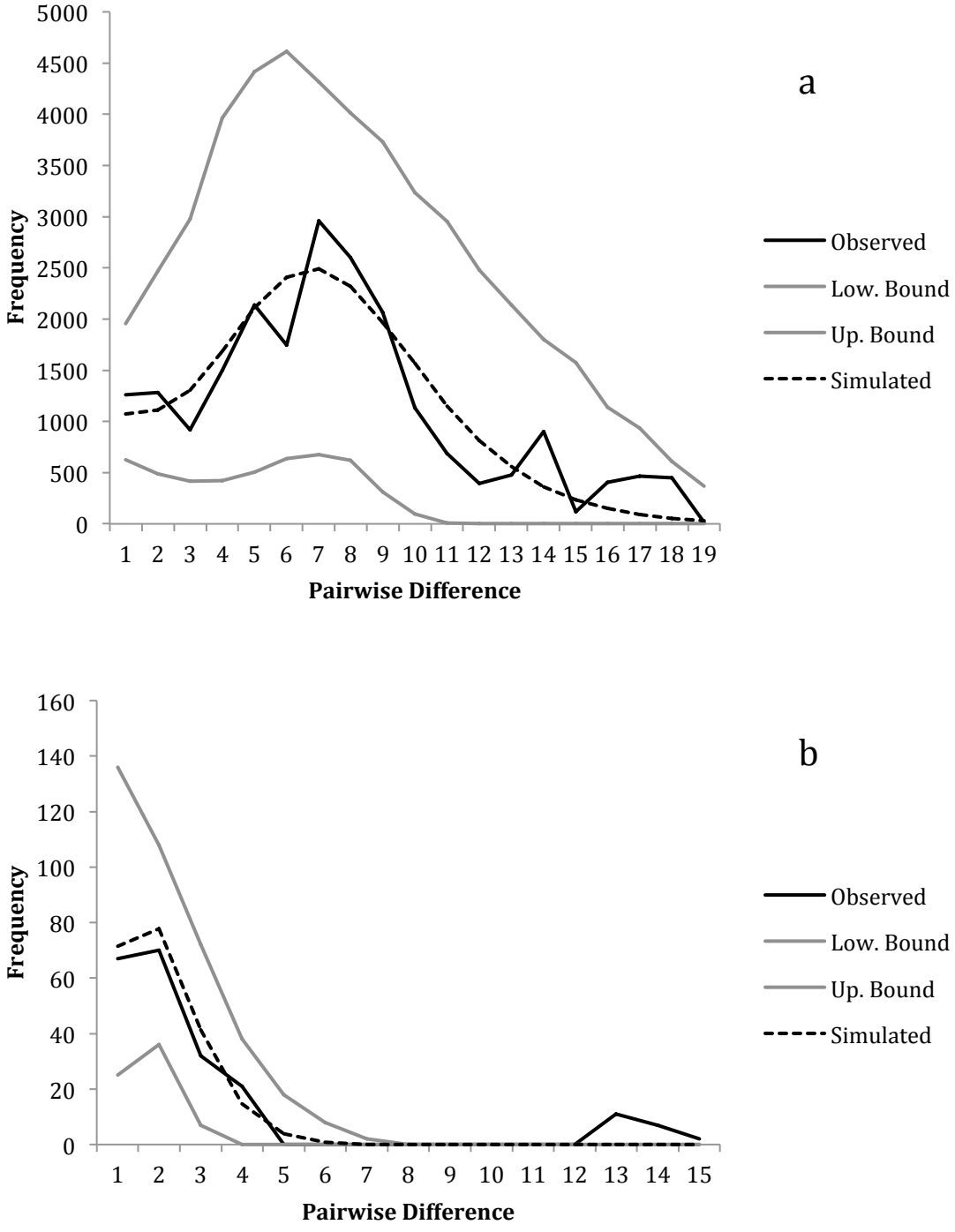

Figure 3.3. Mismatch distribution plot of the observed sequence data (solid black line) and simulated under sudden expansion (dashed black line) based on pairwise differences among contemporary control region haplotypes of a) gray and b) humpback whales. The gray lines represent the 95% confidence interval.

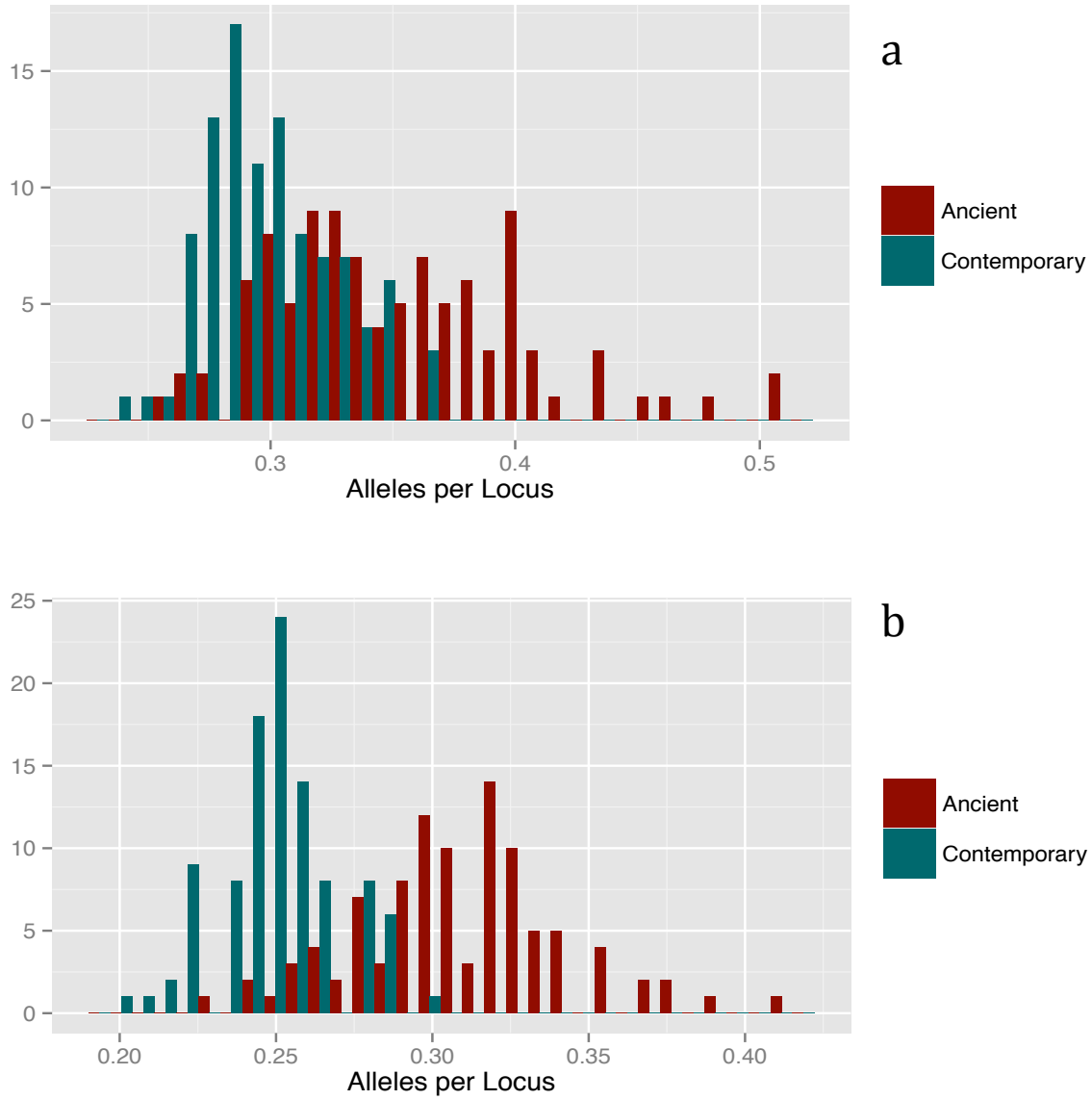

Figure 3.4. Allelic diversity of contemporary and ancient a) gray and b) humpback whales. Proportion of unique alleles at a locus.

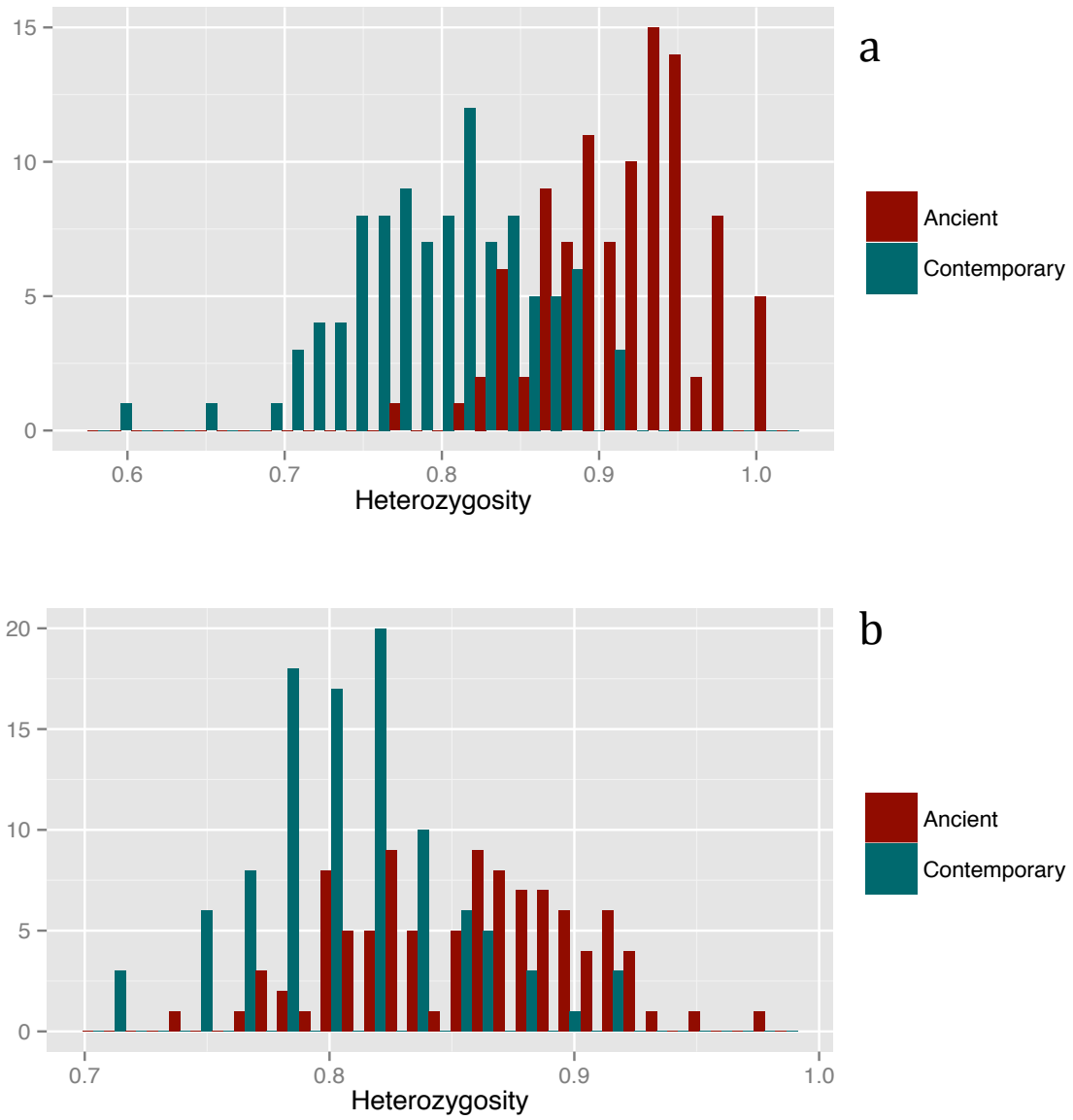

Figure 3.5. Heterozygosity of contemporary and ancient a) gray and b) humpback whales.

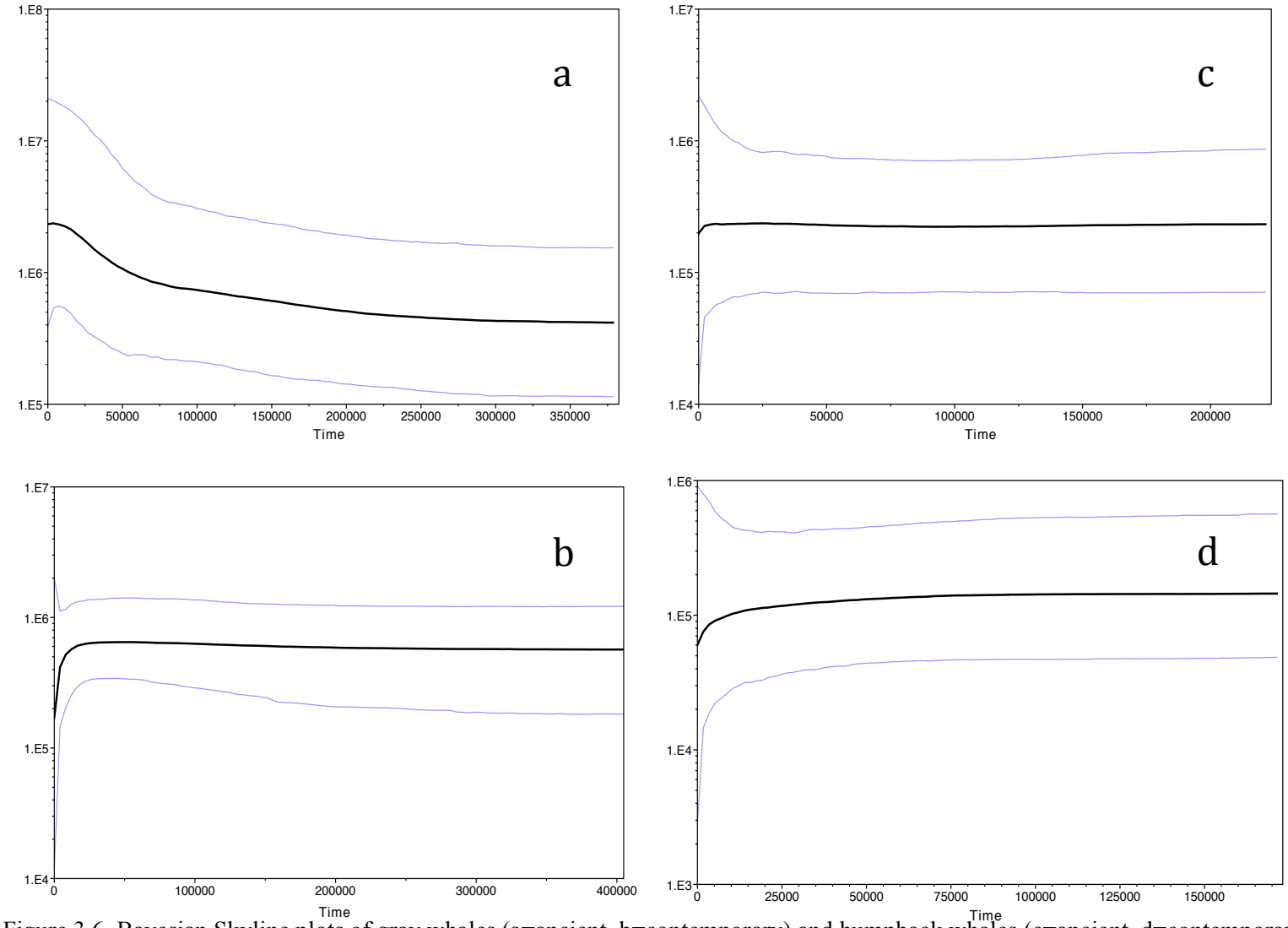

Figure 3.6. Bayesian Skyline plots of gray whales (a=ancient, b=contemporary) and humpback whales (c=ancient, d=contemporary). The y-axis is a logarithmic scale indicating effective female population size multiplied by generation time. Time zero on ancient skyline plots (a,c) indicate pre-whaling, whereas time zero on contemporary skyline plots (b,d) are indicative of present day. Dark lines represent the mean from the estimated genealogies, and blue is the highest posterior density interval, which can be thought of as the confidence interval.

Table 3.1. Summary of the number of ancient and contemporary gray and humpback whales genotyped at the loci previously described by (D'Intino *et al.* 2013).

Locus	Source	Contemporary		Ancient	
		Gray Whales	Humpback Whales	Gray Whales	Humpback Whales
EV94Mn	Valsecchi and Amos (1996)	170	22	N/A	N/A
RW31	Waldick et al. (1999)	176	22	N/A	N/A
SW13	Richard et al. (1996)	158	22	N/A	N/A
GT023	Bérubé et al. (2000)	172	22	N/A	N/A
EV1Pm	Valsecchi and Amos (1996)	174	22	N/A	N/A
TexVet5	Rooney et al. (1999)	173	22	15	18
FCB4	Buchanan et al. (1996)	161	N/A	N/A	N/A
EV14Pm	Valsecchi and Amos (1996)	169	22	N/A	N/A
EV37Mn	Valsecchi and Amos (1996)	162	22	9	21
FCB14	Buchanan et al. (1996)	159	22	N/A	N/A
FCB5	Buchanan et al. (1996)	160	22	N/A	N/A
GATA028	Palsbøll et al. (1997)	166	22	N/A	N/A
FCB17	Buchanan et al. (1996)	165	22	19	20
SW10	Richard et al. (1996)	175	22	N/A	N/A
GATA417	Palsbøll et al. (1997)	172	22	19	21

N/A = Not Applicable

Table 3.2. *M*-ratio results of the contemporary gray and humpback whale microsatellite data. All *M*-ratios are larger than the *Mc*, which indicates a lack of bottleneck detection.

Species	Mutation Rate	Ne	Δg	*ps*	*M*-ratio	*Mc*
Gray Whale	$5x10^{-4}$	15,000	3.5	0.90	0.89047	0.733961
		15,000		0.95	0.89047	0.719347
		35,000		0.90	0.89047	0.775976
		35,000		0.95	0.89047	0.756464
Humpback Whale	$5x10^{-4}$	15,000	3.5	0.90	0.83192	0.612567
		15,000		0.95	0.83192	0.594194
		20,000		0.90	0.83192	0.666482
		20,000		0.95	0.83192	0.644294

Table 3.3. Parameters for BEAST (Bayesian skyline plot)

Parameter								
Time Period	Species	Mutation Rate	Mutation model	Bayesian Skyline (groups)	Priors	MCMC Chains	Burn-Ins	Log Tree every
Ancient	Gray	$3.5x10^{-9}$	TN93 + G (8 cat)	8	Program Defaults	10,000,000	1,000,000	1000
		$4.0x10^{-8}$						
	Humpback	$3.5x10^{-9}$	HKY+I	8	Program Defaults	10,000,000	1,000,000	1000
		$4.0x10^{-8}$						
Contemporary	Gray	$3.5x10^{-9}$	TN93 + G (8 cat)	10	Program Defaults	100,000,000	10,000,000	2000
		$4.0x10^{-8}$						
	Humpback	$3.5x10^{-9}$	HKY	8	Program Defaults	10,000,000	1,000,000	1000
		$4.0x10^{-8}$						

Literature Cited

Alter SE, Newsome SD, Palumbi SR (2012) Pre-whaling genetic diversity and population ecology in eastern Pacific gray whales: Insights from ancient DNA and stable isotopes. *PLoS ONE*, **7**, e35039.

Alter SE, Rynes E, Palumbi SR (2007) DNA evidence for historic population size and past ecosystem impacts of gray whales. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 15162–15167.

Baker CS, Perry A, Bannister JL *et al.* (1993) Abundant mitochondrial DNA variation and world-wide population structure in humpback whales. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 8239–8243.

Baker CS, Steel D, Calambokidis J *et al.* (2013) Strong maternal fidelity and natal philopatry shape genetic structure in North Pacific humpback whales. *Marine Ecology Progress Series*, **494**, 291–306.

Balazs GH, Chaloupka M (2004) Thirty-year recovery trend in the once depleted Hawaiian green sea turtle stock. *Biological Conservation*, **117**, 491–498.

Barnes I, Shapiro B, Lister A *et al.* (2007) Genetic structure and extinction of the woolly wammoth, *Mammuthus primigenius*. *Current Biology*, **17**, 1072–1075.

Bellinger MR, Johnson JA, Toepfer J, Dunn P (2003) Loss of genetic variation in greater prairie chickens following a population bottleneck in Wisconsin, USA. *Conservation Biology*, **17**, 717–724.

Bérubé M, Jorgensen H, McEwing R and Palsbøll PJ (2000) Polymorphic di-nucleotide microsatellite loci isolated from the humpback whale, *Megaptera novaeangliae*. *Molecular Ecology,* **9**, 181–2,83.

Bos DH, Gopurenko D, Williams RN, DeWoody JA (2008) Inferring population history and demography using microsatellites, mitochondrial DNA, and major histocompatibility complex (MHC) genes. *Evolution*, **62**, 1458–1468.

Buchanan FC, Friesen MK, Littlejohn RP and Clayton JW (1996) Microsatellites from the beluga whale *Delphinapterus leucas*. *Molecular Ecology,* **5**, 571–75.

Busch JD, WASER PM, DeWoody JA (2007) Recent demographic bottlenecks are not accompanied by a genetic signature in banner-tailed kangaroo rats (*Dipodomys spectabilis*). *Molecular Ecology*, **16**, 2450–2462.

Carroll EL, Jackson JA, Paton D, Smith TD (2014) Two intense decades of 19th century whaling precipitated rapid decline of right whales around New Zealand and East Australia. *PLoS ONE*, **9**, e93789.

Chan YL, Anderson CNK, Hadly EA (2005) Bayesian estimation of the timing and severity of a population bottleneck from ancient DNA. *PLoS Genetics*, **4**, e59.

Chen M-H, Kuo L, Lewis PO (2013) *Bayesian Phylogenetics: Methods, Computational Algorithms, and Applications*. C R C Press LLC.

Clayton D (2000) *Islands of Truth: The Imperial Fashioning of Vancouver Island*. UBC Press, Vancouver.

Cooper A, Poinar HN (2000) Ancient DNA: do it right or not at all. *Science*, **289**, 1139–1139.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.
Coté C (2010) *Spirits of our Whaling Ancestors: Revitalizing Makah & Nuu-chah-nulth Traditions.* UBC Press, Vancouver.
COSEWIC (2011) COSEWIC assessment and status report on the Humpback Whale (*Megaptera novaeangliae*) in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. x + 32 pp.
Cypriano-Souza AL, Fernandez GP, Lima-Rosa CAV, Engel MH, Bonatto SL (2010) Microsatellite Genetic Characterization of the Humpback Whale (*Megaptera novaeangliae*) Breeding Ground off Brazil (Breeding Stock A). *Journal of Heredity*, **101**, 189–200.
D'Intino A, Darling JD, Urbán J, Frasier TR (2013) Lack of nuclear differentiation suggests reproductive connectivity between the "southern feeding group" and the larger population of eastern North Pacific gray whales, despite previous detection of mitochondrial differences. *Journal of Cetacean Resource Managment*, **13**, 97–104.
Drummond AJ (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22**, 1185–1192.
Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.
Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.
Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.
Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics online*, **1**, 47–50.
Foote AD, Hofreiter M, Morin PA (2012) Ancient DNA from marine mammals: Studying long-lived species over ecological and evolutionary timescales. *Annals of Anatomy - Anatomischer Anzeiger*, **194**, 112–120.
Frasier TR, Koroscil SM, White BN, Darling JD (2011) Assessment of population substructure in relation to summer feeding ground use in the eastern North Pacific gray whale. *Endangered Species Research*, **14**, 39–48.
Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.
Fu YX, Li W-H (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.
Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305–318.
Griekspoor A and Groothuis T (2006) 4Peaks Version 1.7.2. nucleobytes.com
Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591.
Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. *Current Anthropology*, **34**, 483–496.
Ho SYW, Shapiro B (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources*, **11**, 423–434.

Hu Y, Guo YU, Qi D *et al.* (2011) Genetic structuring and recent demographic history of red pandas (Ailurus fulgens) inferred from microsatellite and mitochondrial DNA. *Molecular Ecology*, **20**, 2662–2675.
Johnson JA, Dunn PO, Bouzat JL (2007) Effects of recent population bottlenecks on reconstructing the demographic history of prairie-chickens. *Molecular Ecology*, **16**, 2203–2222.
Keane TM, Creevey CJ, Pentony MM, Naughton TJ, Mclnerney JO (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evolutionary Biology*, **6**, 29.
Kelley K, Preacher KJ (2012) On effect size. *Psychological Methods*, **17**, 137–152.
Koringa PG, Joshi CG, Solanki JV, Rank DN (2008) Genetic characterization and bottleneck studies in Kathiawari horse breed of India. *Haryana Vet*, **47**, 77–83.
Lambertsen RH (1987) A biopsy system for large whales and its use for cytogenetics. *Journal of Mammalogy* **68**, 443-445.
Larkin MA, Blackshields G, Brown NP *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
Le Page SL, Livermore RA, Cooper DW, Taylor AC (2001) Genetic analysis of a documented population bottleneck: introduced Bennett's wallabies (Macropus rufogriseus rufogriseus) in New Zealand. *Molecular Ecology*, **9**, 753–763.
Lessios HA, Garrido MJ, Kessing BD (2001) Demographic history of *Diadema antillarum*, a keystone herbivore on Caribbean reefs. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 2347–2353.
Lorenzen ED, Nogués-Bravo D, Orlando L *et al.* (2011) Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature*, **479**, 359–364.
Luikart G, Cornuet JM (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology*, **12**, 228–237.
Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89**, 238–247.
Martínez Barrio Á, Lagercrantz E, Sperber GO, Blomberg J, Bongcam-Rudloff E (2009) Annotation and visualization of endogenous retroviral sequences using the Distributed Annotation System (DAS) and eBioX. *BMC Bioinformatics*, **10**, S18.
McLeod BA, Brown MW, Frasier TR, White BN (2010) DNA profile of a sixteenth century western North Atlantic right whale (Eubalaena glacialis). *Conservation Genetics*, **11**, 339–345.
McLeod BA, Brown MW, Moore MJ *et al.* (2008) Bowhead whales, and not right whales, were the primary target of 16th-to 17th-century Basque whalers in the western North Atlantic. *Arctic*, 61–75.
McNeely R, Dyke AS, Southon JR (2006) Canadian marine reservoir ages, preliminary assessment. *Geological Survey of Canada Open File 5049*.
Moore MJ (2014) How we all kill whales. *ICES Journal of Marine Science*, **71**, 760-763.
Mulligan CJ (2006) Anthropological applications of ancient DNA: Problems and prospects. *American Antiquity*, 365–380.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1–10.
Nichol LM, Gregr EJ, Flinn R *et al.* (2002) *British Columbia Commercial Whaling Catch Data 1908 to 1967 : A Detailed Description of the B.C. Historical Whaling Database*. Fisheries and Oceans Canada, Nanaimo.
Nyström V, Humphrey J, Skoglund P *et al.* (2012) Microsatellite genotyping reveals end-Pleistocene decline in mammoth autosomal genetic variation. *Molecular Ecology*, **21**, 3391–3402.
Palsbøll PJ, Bérubé M, Larsen AH and Jorgensen H (1997) Primers for the amplification of tri- and tetramer microsatellite loci in baleen whales. *Molecular Ecology*, **6**, 893–95.
Palsbøll PJ, Larsen F and Hansen ES (1991) Sampling of skin biopsies from free-ranging large cetaceans in west Greenland: development of new biopsy tips and bolt designs. *Reports of the International Whaling Commission* (Special Issue 13):71-79.
Palsbøll PJ, Zachariah Peery M, Olsen MT, Beissinger SR, Bérubé M (2012) Inferring recent historic abundance from current genetic diversity. *Molecular Ecology*, **22**, 22–40.
Pauly D (1995) Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology & Evolution*, **10**, 430.
Peery MZ, Kirby R, Reid BN *et al.* (2012) Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular Ecology*, **21**, 3403–3418.
Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity*, **90**, 502–503.
Prost S, Anderson CNK (2011) TempNet: a method to display statistical parsimony networks for heterochronous DNA sequence data. *Methods in Ecology and Evolution*, **2**, 663–667.
R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org
Ramakrishnan U, Hadly EA, Mountain JL (2005) Detecting past population bottlenecks using temporal genetic data. *Molecular Ecology*, **14**, 2915–2922.
Rambaut A, Drummond AJ (2007) Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer
Rastogi T, Brown MW, McLeod BA *et al.* (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**, 1647–1654.
Reeves RR (2002) The origins and character of "aboriginal subsistence"whaling: a global review. *Mammal Review*, **32**, 71–106.
Richard KR, Whitehead H and Wright JM (1996) Polymorphic microsatellites from sperm whales and their use in the genetic identification of individuals from naturally sloughed pieces of skin. *Molecular Ecology,* **5**, 313–15.
Robert L, Brownell J, Yablokov AV (2009) Illegal and pirate whaling. In: *Encyclopedia of Marine Mammals* (eds Perrin WF, Wursig B, Thewissen JGM), p. 1352. Academic Press, Burlington, MA, USA.
Rogers AR, Harpending HC (1992) Population growth makes waves in the distribution of

pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.
Rooney AP, Honeycutt RL, Derr JN (2001) Historical population size change of bowhead whales inferred from DNA sequence polymorphism data. *Evolution*, **55**, 1678–1685.
Rooney AP, Merritt DB and Derr JN (1999) Microsatellite diversity in captive bottlenose dolphins (*Tursiops truncatus*). *Journal of Heredity*, **90**, 228–31.
Rose NA, Janiger D, Parsons ECM, Stachowitsch M (2011) Shifting baselines in scientific publications A case study using cetacean research. *Marine Policy*, **35**, 477–482.
Sastre N, Vilà C, Salinas M *et al.* (2010) Signatures of demographic bottlenecks in European wolf populations. *Conservation Genetics*, **12**, 701–712.
Schmidt HA, Strimmer K, Vingron M, Haeseler von A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*, **18**, 502–504.
Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, **69**, 82–90.
Sheppard C (1995) The Shifting Baseline Syndrome. *Marine Pollution Bulletin*, **30**, 766–767.
Smith BT, Escalante P, Baños BEH *et al.* (2011) The role of historical and contemporary processes on phylogeographic structure and genetic diversity in the Northern cardinal, *Cardinalis cardinalis*. *BMC Evolutionary Biology*, **11**, 136.
Smith TD, Reeves RR (2003) Estimating American 19th Century Catches of Humpback Whales in the West Indies and Cape Verde Islands. *Caribbean Journal of Science*, **39**, 286–297.
Stiller M, Baryshnikov G, Bocherens H *et al.* (2010) Withering away--25,000 years of genetic decline preceded cave bear extinction. *Molecular Biology and Evolution*, **27**, 975–978.
Taberlet P, Griffin S, Goossens B *et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic acids research*, **24**, 3189–3194.
Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
Taylor BL, Chivers SJ, Larese J, Perrin WF (2007) Generation length and percent mature estimates for IUCN assessments of cetaceans. *Southwest Fisheries Science Center*.
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic acids research*, **25**, 4876–4882.
Torres-Florez JP, Hucke-Gaete R, Rosenbaum H, Figueroa CC (2014) High genetic diversity in a small population: the case of Chilean blue whales. *Ecology and Evolution*, **4**, 1398–1412.
Valsecchi E and Amos W (1996) Microsatellite markers for the study of cetacean populations. *Molecular Ecology* **5**, 151–56.
Villesen P (2007) FaBox: an online toolbox for fasta sequences. *Molecular Ecology Notes*, **7**, 965–968.
Waldick RC, Kraus S, Brown MW, White BN (2002) Evaluating the effects of historic bottleneck events: an assessment of microsatellite variability in the endangered, North Atlantic right whale. *Molecular Ecology*, **11**, 2241–2249.

Wang JY, Frasier TR, Yang SC, White BN (2008) Detecting recent speciation events: the case of the finless porpoise (genus Neophocaena). *Heredity*, **101**, 145–155.
Werle E, Schneider C, Renner M, Völker M, Fiehn W (1994) Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucleic acids research*, **22**, 4354.
Wilcoxon F (1945) Individual comparisons by ranking methods. *Biometrics bulletin*, **1**, 80–83.
Yang DY, Watt K (2005) Contamination controls when preparing archaeological remains for ancient DNA analysis. *Journal of Archaeological Science*, **32**, 331–336.
Yoshida H, Yoshioka M, Shirakihara M, Chow S (2001) Population structure of finless porpoises (Neophocaena phocaenoides) in coastal waters of Japan based on mitochondrial DNA sequences. *Journal of Mammalogy*, **82**, 123–130.

Chapter 4

General Discussion

Genetic Methods and Species Identification

Genetic methods have become useful in more fields than just population genetics, as anthropology and conservation management can also benefit from these techniques and advancements in science. Particularly, genetic species identification played a key role in achieving the objectives of this project. The initial step of conducting species identification via genetic methods was necessary for 3 reasons: 1) to demonstrate that ancient bone fragments, that would have been left unidentified through conventional comparative morphological methods, could in fact be assigned to a species; 2) to identify that gray, humpback, right and fin whales were once hunted at these traditional First Nations whaling sites on Vancouver Island; and 3) to obtain a pre-commercial whaling DNA sample set for Eastern North Pacific gray whales and North Pacific humpback whales.

The first reason was important because genetic methods could potentially have an affect on how species identification will be conducted for future archaeological projects in the area. Genetic methods are already slowly being incorporated into current archaeological work (Newman *et al.* 2002; Nicholls *et al.* 2003; Yang *et al.* 2005; Grier *et al.* 2013). It could also mean that past species identifications that were done via comparative morphology can be confirmed or altered through the use of genetic identification, if the DNA is not too degraded.

The second reason the initial species identification is of particular importance is for the investigation of our First Nations culture on the west coast of Vancouver Island. Most of the history in this area is kept within oral history that is passed down from one

generation to the next (Harkin 1998; Coté 2010). With the disconnection of many First Nations to their whaling traditions, there is an increased potential for the loss or miscommunication of this oral history (Coté 2010). The oral history stated that gray and humpbacks were two of the main targets of the Nuu-chah-nulth First Nations, and the results of the species identification of the sample site on Echachist Island and surround archaeological sites lend support to this (Clayton 2000; Monks *et al.* 2001; McMillan & Claire 2005; Coté 2010). The importance of the continuation of this oral history in First Nations cultures is recognizable, and the support from scientific research can assist in ensuring that there is no loss of key information. This is especially critical with the resumption of many First Nations whaling traditions, such as the Maa-nulth's treaty that includes the right to begin whaling and will lead to the resumption of many traditions that were associated with this practice (Coté 2010).

The third and final reason why the initial species identification of the whales bones found at these traditional whaling sites was necessary, was for developing a pre-commercial whaling genetic baseline for gray and humpback whale populations that utilize the coastal waters off Vancouver Island as a summer feeding ground (Frasier *et al.* 2011). This pre-commercial baseline is key for investigating the effects of anthropological caused events, such as the mass depletion of whales in our oceans, have had on the genetic variation of these species. It was seen though the use of current and popular tests that genetic signatures of reduction within the gray and humpback whale's nuclear and mitochondrial DNA were not detectable, as not enough time has passed since the recent bottleneck event of commercial whaling. The genetics are not following the

predictable signatures that would be seen in a species recovering from a population decrease, again a demonstration a limitation when there has not been enough passage of time for those species that have a longer life-span and generation time. Genetic signatures demonstrating difference between the pre- and post-whaling sample set were detectable when these two time periods were compared. This use of multiple time points in inferring demography is becoming more popular. Through the investigation of the pattern of demographic change over time, by utilizing two time periods, we can even see that commercial whaling may have erased historical genetic signatures, which could lead to phenomena known as "shifting baselines" and then result in inaccurate management goals (Pauly 1995; Sheppard 1995). Biologists are well aware that commercial whaling depleted whale populations, but realizing that we do not have accurate pre-exploitation sizes for these species means that we do not know, quantitatively, to what severity this depletion was. This reduction was seen to have a negative affect on the genetic variability of gray and humpback whales, but we are still unaware to what impact this will have on their future, in regards to reproductive success and population growth, and it should be a cause for concern and we should be more conservative in setting catch limits and management goals.

There is no doubt that genetic methods can assist in the field of conservation biology, but it is also proving to be very helpful in many other areas of research such as archaeology, cultural anthropology and history (Mulligan 2006; Cai *et al.* 2007; Rick & Lockwood 2013). In some cases it is able to lend support to previous theories and studies, such as discovering which species were once taken by our ancestors and what

affect the early humans had on the extinction of megafauna of the Pleistocene Age (McLeod *et al.* 2008; Stiller *et al.* 2010; Nyström *et al.* 2012; Allentoft *et al.* 2014). It can also disprove previous thought and lead to a new direction of research of why species populations are endangered or not recovering in the way scientist would anticipate (Rastogi *et al.* 2004; McLeod *et al.* 2008).

Future Direction

If this project was to continue and expand, it would be ideal to add four key elements: 1) increase sampling sites along the coast of Vancouver Island, as this would assist in answering the question as to whether or not there is a geographical difference in the species targeted by First Nations groups, 2) using genetic species identification techniques to confirm the species of those bones what were identified by osteological methods at both the Barkley Sound and Ozette archaeological sites, 3) conduct genetic species identification on those bones from these archaeological sites that were previously left unidentified due to the limit of fragmentation, and 4) increasing the number of samples that were radio-carbon dated. Getting a more precise age of the bones recovered could led to dividing the ancient time period into two or more. This could assist in understanding what affect historic versus pre-historic environmental changes has on these whale species.

As the cost of conducting genetic research slowly decreases and the power of computers advance, we can expect to see an increase in use of genetic methods to answer many questions. These genetic data cannot only provide information about the history of

a species but can also be helpful for predicting its future. Particularly for the future of this project, utilizing historic and pre-historic ocean temperature data may provide some indication on what effect past environmental changes had on the whale populations, as similar work has been done on now extinct terrestrial species populations (Nyström *et al.* 2012). These data could explain pre-historic population increases or decreases and aid in creating models to predict future changes to demography given the issue of global warming (Willis & MacDonald 2011). Climate change modeling is already being used to predict habitat change in extant species, such as polar bears (Hunter *et al.* 2010). In order to properly manage and maintain the diversity of life found on land and in the oceans, researchers must utilize every means available, with genetics becoming one of the key tools.

Literature Cited

Allentoft ME, Heller R, Oskam CL *et al.* (2014) Extinct New Zealand megafauna were not in decline before human colonization. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 4922–4927.

Cai D-W, Han L, Zhang X-L, Zhou H, Zhu H (2007) DNA analysis of archaeological sheep remains from China. *Journal of Archaeological Science*, **34**, 1347–1355.

Clayton D (2000) *Islands of Truth: The Imperial Fashioning of Vancouver Island*. UBC Press, Vancouver.

Coté C (2010) *Spirits of our Whaling Ancestors: Revitalizing Makah & Nuu-chah-nulth Traditions.* UBC Press, Vancouver.

Frasier TR, Koroscil SM, White BN, Darling JD (2011) Assessment of population substructure in relation to summer feeding ground use in the eastern North Pacific gray whale. *Endangered Species Research*, **14**, 39–48.

Grier C, Flanigan K, Winters M *et al.* (2013) Using ancient DNA identification and osteometric measures of archaeological Pacific salmon vertebrae for reconstructing salmon fisheries and site seasonality at Dionisio Point, British Columbia. *Journal of Archaeological Science*, **40**, 544–555.

Harkin M (1998) Whales, chiefs, and giants: an exploration into Nuu-chah-nulth political thought. *Ethnology*, 317–332.

Hunter CM, Caswell H, Runge MC *et al.* (2010) Climate change threatens polar bear populations: a stochastic demographic analysis. *Ecology*, **91**, 2883–2897.

McLeod BA, Brown MW, Moore MJ *et al.* (2008) Bowhead whales, and not right whales, were the primary target of 16th-to 17th-century Basque whalers in the western North Atlantic. *Arctic*, 61–75.

McMillan AD, Claire DES (2005) *Ts'ishaa*. Archaeology Press. Department of Archaeology, Simon Fraser University., Burnaby.

Monks GG, McMillan AD, St Claire DE (2001) Nuu-chah-nulth whaling: archaeological insights into antiquity, species preferences, and cultural importance. *Arctic anthropology*, 60–81.

Mulligan CJ (2006) Anthropological applications of ancient DNA: Problems and prospects. *American Antiquity*, 365–380.

Newman ME, Parboosingh JS, Bridge PJ, Ceri H (2002) Identification of archaeological animal bone by PCR/DNA analysis. *Journal of Archaeological Science*, **29**, 77–84.

Nicholls A, Matisoo Smith E, Allen MS (2003) A novel application of molecular techniques to Pacific archaeofish remains. *Archaeometry*, **45**, 133–147.

Nyström V, Humphrey J, Skoglund P *et al.* (2012) Microsatellite genotyping reveals end-Pleistocene decline in mammoth autosomal genetic variation. *Molecular Ecology*, **21**, 3391–3402.

Pauly D (1995) Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology & Evolution*, **10**, 430.

Rastogi T, Brown MW, McLeod BA *et al.* (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**, 1647–1654.

Rick TC, Lockwood R (2013) Integrating paleobiology, archeology, and history to inform

biological conservation. *Conservation Biology*, **27**, 45–54.
Sheppard C (1995) The shifting baseline syndrome. *Marine Pollution Bulletin*, **30**, 766–767.
Stiller M, Baryshnikov G, Bocherens H *et al.* (2010) Withering away--25,000 years of genetic decline preceded cave bear extinction. *Molecular Biology and Evolution*, **27**, 975–978.
Willis KJ, MacDonald GM (2011) Long-term ecological records and their relevance to climate change predictions for a warmer world. *Annual Review of Ecology, Evolution, and Systematics*, **42**, 267–287.
Yang DY, Woiderski JR, Driver JC (2005) DNA analysis of archaeological rabbit remains from the American Southwest. *Journal of Archaeological Science*, **32**, 567–578.