DOCTOR OF PHILOSOPHY

Evolution & Extinction of the Great Auk
A Palaeogenomic Approach

Thomas, Jessica

Award date:
2018

Awarding institution:
Bangor University
University of Copenhagen

PRIFYSGOL
BANGOR
UNIVERSITY

KØBENHAVNS UNIVERSITET
NATURAL HISTORY MUSEUM OF DENMARK

Evolution & Extinction of the Great Auk: A Palaeogenomic Approach

A thesis submitted for the degree of Doctor of Philosophy at

Bangor University

&

University of Copenhagen

Double Degree PhD

Jessica Emma Thomas

2018

Declaration and Consent

Details of the Work

I hereby agree to deposit the following item in the digital repository maintained by Bangor University and/or in any other repository authorized for use by Bangor University.

Author Name: Jessica Emma Thomas

Title: Evolution & Extinction of the Great Auk: A Palaeogenomic Approach

Supervisor/Department: Dr Michael Knapp, Otago University; Professor Gary Carvalho, Biological Sciences, Bangor University; Professor M. Thomas Gilbert, University of Copenhagen, Dr John Stewart, University of Bournemouth; Professor Michael Hofreiter, University of Potsdam.

Funding body (if any): NERC & Natural History Museum of Denmark

Qualification/Degree obtained: PhD Biological Science

This item is a product of my own research endeavours and is covered by the agreement below in which the item is referred to as "the Work". It is identical in content to that deposited in the Library, subject to point 4 below.

Non-exclusive Rights

Rights granted to the digital repository through this agreement are entirely non-exclusive. I am free to publish the Work in its present version or future versions elsewhere.

I agree that Bangor University may electronically store, copy or translate the Work to any approved medium or format for the purpose of future preservation and accessibility. Bangor University is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.

Bangor University Digital Repository

I understand that work deposited in the digital repository will be accessible to a wide variety of people and institutions, including automated agents and search engines via the World Wide Web.

I understand that once the Work is deposited, the item and its metadata may be incorporated into public access catalogues or services, national databases of electronic theses and dissertations such as the British Library's EThOS or any service provided by the National Library of Wales.

Statement 1:

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless as agreed by the University for approved dual awards.

Signed …….. (candidate)

Date ……..

Statement 2:

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

All other sources are acknowledged by footnotes and/or a bibliography.

Signed ……. (candidate)

Date …….

Statement 3:

I hereby give consent for my thesis, if accepted, to be available for photocopying, for inter-library loan and for electronic repositories, and for the title and summary to be made available to outside organisations.

Signed ……. (candidate)

Date …….

NB: Candidates on whose behalf a bar on access has been approved by the Academic Registry should use the following version of **Statement 3:**

Statement 3 (bar):

I hereby give consent for my thesis, if accepted, to be available for photocopying, for inter-library loans and for electronic repositories after expiry of a bar on access.

Signed …… (candidate)

Date ……

Statement 4:

Choose **one** of the following options

a) I agree to deposit an electronic copy of my thesis (the Work) in the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University and where necessary have gained the required permissions for the use of third party material.	

b) I agree to deposit an electronic copy of my thesis (the Work) in the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University when the approved **bar on access** has been lifted.	
c) I agree to submit my thesis (the Work) electronically via Bangor University's e-submission system, however I **opt-out** of the electronic deposit to the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University, due to lack of permissions for use of third party material.	

Options B should only be used if a bar on access has been approved by the University.

In addition to the above I also agree to the following:

1. That I am the author or have the authority of the author(s) to make this agreement and do hereby give Bangor University the right to make available the Work in the way described above.

2. That the electronic copy of the Work deposited in the digital repository and covered by this agreement, is identical in content to the paper copy of the Work deposited in the Bangor University Library, subject to point 4 below.

3. That I have exercised reasonable care to ensure that the Work is original and, to the best of my knowledge, does not breach any laws – including those relating to defamation, libel and copyright.

4. That I have, in instances where the intellectual property of other authors or copyright holders is included in the Work, and where appropriate, gained explicit permission for the inclusion of that material in the Work, and in the electronic form of the Work as accessed through the open access digital repository, *or* that I have identified and removed that material for which adequate and appropriate permission has not been obtained and which will be inaccessible via the digital repository.

5. That Bangor University does not hold any obligation to take legal action on behalf of the Depositor, or other rights holders, in the event of a breach of intellectual property rights, or any other right, in the material deposited.

6. That I will indemnify and keep indemnified Bangor University and the National Library of Wales from and against any loss, liability, claim or damage, including without limitation any related legal fees and court costs (on a full indemnity bases), related to any breach by myself of any term of this agreement.

Signature: ……

Date: …….

Abstract

The great auk, (*Pinguinus impennis*), was a flightless bird, once abundant and widely distributed across the North Atlantic. It was, however, heavily exploited for its eggs, meat, oil and feathers, and latterly as a display item. Despite wide scientific and general interest in the species, it remains unclear if hunting alone was responsible for its demise, or whether it was already in decline due to other factors such as climate-driven environmental change. Here, we address various issues relating to the great auks' evolution and extinction, including morphometric differentiation and population genetics. In contrast to previous findings, morphometric studies on humerus samples indicated no-significant population-specific size variation. High-throughput sequencing of ancient DNA (aDNA) samples of complete mitochondrial genomes from 41 great auks was undertaken from samples across the Holocene and Late Pleistocene range. Data showed consistently significant high levels of genetic diversity and gene flow persisting through time and across their range. Demographic reconstructions revealed the great auk had a large and stable effective population size, with no evidence of decline, associated with periods of climatic change. Population viability analysis indicated that harvesting of 5-7% of the total population would be required to cause extinction in a period of fewer than 350 years; levels commensurate with documented observations. Findings are consistent with the current consensus that human hunting was the primary cause of the great auks' extinction. Additional analysis of nuclear DNA yielded no data due to low coverage of the 495 targeted markers. Related analyses on mitochondrial genomes of five candidate specimens and the organs from the last documented pair of great auks allowed for a 170-year-old mystery to be solved, by matching the male organs with the skin currently on display at the RBINS, Brussels. Collectively, our findings yield insights into the lives of an iconic extinct bird.

Abstract in Danish

Resume

Gejrfuglen, (*Pinguinus impennis*), var en ikke-flyvende fugl, som engang var vidt udbredt i Nordatlanten i store antal. Den blev imidlertid voldsomt udnyttet for sine æg, kød, olie og fjer og senere som udstillingseksemplar. Til trods for en bred videnskabelig og generel interesse i denne art, er det forblevet uklart hvorvidt jagt alene var skyld i dens udryddelse, eller om arten allerede var i nedgang på grund af andre faktorer så som klima-drevne miljøændringer. I dette studie adresserer vi diverse emner relateret til gejrfuglens evolution og udryddelse, heriblandt morfometrisk differentiering og populationsgenetik. I modsætning til tidligere studier, viste morfometriske studier af humerus-prøver en ikke-signifikant populationsspecifik-størrelsesvariation. High-throughput sekventering af ancient-DNA (aDNA)-prøver fra komplette mitokondrielle genomer fra 41 gejrfugle blev udført med prøver fra hele den Holocæne og Sen Pleistocæne periode. Data viste konsekvent høje signifikante niveauer af genetisk diversitet og gen-flow, som var konsistent over tid og udbredelse. Demografiske rekonstruktioner afslørede at gejrfuglen havde en stor og stabil effektiv populationsstørrelse uden nogen beviser for nedgang relateret til perioder med klimaforandringer. Analyser af populationslevedygtighed indikerede at jagt på 5-7% af den totale population ville være nødvendig for at føre til artens udryddelse i løbet af en periode på under 350 år; hvilket er en rimelig periode ifølge dokumenterede observationer. Resultaterne passer med den nuværende konsensus at menneskets jagt var den primære årsag til gejrfuglens udryddelse. Øvrige analyser af nuklear DNA gav ingen data på grund af lav dækning af de 495 målrettede markører. Relaterede analyser på mitokondrielle genomer af fem udvalgte eksemplarer og organer fra det sidste dokumenterede gejrfuglepar gav mulighed for at løse et 170 år gammelt mysterium, ved at matche hannens organer med det skind som i øjeblikket fremvises på RBINS, i Bruxelles. Vores fund kaster tilsammen nyt lys over en ikonisk udryddet fugls liv.

Acknowledgements

It is with immense gratitude that I acknowledge the people who have made this project possible.

I am very grateful to the funders of this project, Natural Environment Research Council (NERC)(NE/L501694/1) for providing the funding for this research, the Natural History Museum of Denmark for financial support, and the European Society for Evolutionary Biology (ESEB) and Genetics Society, for awarding travel grants. To Bangor University, as my host institution and the University of Copenhagen, my partner institution, and Otago University who hosted me during my time in New Zealand. It was a pleasure and a privilege to work at these world-renowned institutes.

I extend my sincere thanks to the numerous institutions that provided samples to the project. Without you this project really wouldn't have been possible. Thank you for willingness to be part of this study, the thought-provoking conversations and continued support and interest in the project. I would also like to thank all the co-authors of the papers for their input and valuable feedback, in particular, Errol Fuller and Nic Rawlence. Thanks to James and Nathan, for the help and support with lab work, and Filipe, Mike and Sama, for all the bioinformatic help.

I was fortunate enough to have a great team of supervisors who have all supported and motivated me throughout. Michi, thank you for the support, and the helpful and constructive feedback on manuscripts and thesis chapters. John, thanks for all the help with the morphometric work, it has been enormously helpful, and for providing support, enthusiasm and feedback. Tom, thank you for giving me the fantastic opportunity to work within your group. Without a doubt, this experience has been invaluable to the project and we would not have been able to achieve what we have, had it not been for the chance to work within your group. Thank you for the support and input into the project throughout (and your ability to provide feedback in record breaking time!). You and the rest of the EvoGenomics group were welcoming from the start, and my stay in Copenhagen has given me some of the best times of my life. Gary, thank you for always being there to offer support and interesting conversation. I am very grateful to you for you input into the whole project and for the continued support and regular skype meetings when I jetted off around the world! It has been a pleasure to work with you and the rest of MEFGL. Finally, Michael, you have been the best supervisor that I ever could have hoped for and I am grateful beyond words for everything you have done. Even

from 12,000 miles away you have always been there via skype. Your continued support, help, and input, at all stages of the project has been instrumental to what has been achieved. Thank you!

The nomadic lifestyle I led during my PhD has allowed me to meet and work with some of the greatest people I could ever have hoped to meet. I have been fortunate to work with and learn from, some of the best in the field. I thank all members of MEFGL in Bangor, EvoGenomics in Copenhagen and Bioanthropology group in Dunedin. Each group has bought with it a host of interesting people and it was a pleasure to meet you all. A special thanks to my PhD BFF, Alice Evans. My roomie, Bangor Babe, travel companion and all-round motivator, supplier of tea and cat memes. Thank you for being there through it all and helping me to get to the end. Meow.

A huge thank you to my family and friends, without whom I would not have made it this far. I am so grateful to each and every one of you for the continued support, patience and encouragement. A special thanks to Jen, Sarah and Sarah for believing in me and pushing me to the end. Mum, Dad, Beci and Adam, thank you for always being my biggest supporters and for putting up with me.

Finally, to Joe. Thank you for your unwavering love and support. Thank you for keeping the faith, pushing me to achieve, supporting me, but most of all for being you.

Contents

List of Figures

List of Tables

Chapter 1: General Introduction

1.1 Extinction

Over 99% of species estimated to have once existed are now extinct (Barnosky *et al.*, 2011). Extinction is frequently described as a natural part of evolution (Halliday, 1978; Olson, 1989; Barnosky *et al.*, 2011), and thought to be beneficial to evolution as it removes species which are not well adapted for survival (Raup, 1981). Nonetheless, preventing species loss is one of the main driving forces behind conservation practices (Olson, 1989).

Understanding the mechanisms of species extinction is of great importance to conservation biologists (Owens and Bennett, 2000; Purvis *et al.*, 2000). Various methods have been used to examine and determine extinction risk, and several models/frameworks have been designed to help identify species/populations at risk of extinction (Sekercioglu *et al.*, 2008; McKinney, 1997; Thomas *et al.*, 2004; Pimm *et al.*, 2006; Jetz *et al.*, 2007; Sekercioglu *et al.*, 2012; Machado & Loyola, 2013). The cause of species extinction is highly debated and certain traits are said to make a species susceptible to extinction, such as specialised habitat preference, endemism to island, and large body size (McKinney, 1997), but the two most commonly cited causes are humans and climate change, or a combination of both (Shapiro *et al.*, 2004; Wroe *et al.*, 2006; Campos *et al.*, 2010; Prost *et al.*, 2010). Humans are known to impact species and ecosystems through habitat disturbance, introduction of species, and hunting. While for several extinctions, the role humans have played is widely accepted, it is agreed generally that we cannot overlook climatic influences at a time of major environmental change (Wroe *et al.*, 2006).

It has been noted in several studies that understanding how climate change affects species and how they respond to it is extremely important for future conservation (Hadly *et al.*, 2004; Leonard, 2008; MacDonald *et al.*, 2008; O'Corry-Crowe, 2008; de Bruyn *et al.*, 2009; Prost *et al.*, 2010). During periods of climatic change, it is thought that species can respond in three ways; species can track changing habitats, adapt to the environmental change or become extinct (de Bruyn *et al.*, 2011). The study of extinction can yield valuable insights into the drivers and dynamics of global biodiversity, and therefore there is value in examining the lives of species that have become extinct (Halliday, 1978). One way in which we can gain an insight into better understanding the processes associated with why some species have gone extinct while others have survived is to look at the genetics of extinct species. As we are now in an age where molecular techniques are widely used to address many ecological and evolutionary questions,

it follows that examining a species' genetic diversity and population genetic structure can help address some of the questions surrounding loss of biodiversity and population/species extinctions (Bellard *et al.*, 2012). Ancient DNA (aDNA), offers an especially informative approach to explore features such as past trends in population connectivity, size, geographic distribution, levels of genetic diversity and the response to environmental change, hunting or species introductions. Over the past decade, aDNA has been used to examine species demographic histories and investigate the relative impact and contribution of humans/environmental factors to be assessed in a way not previously possible (Leonard, 2008; de Bruyn *et al.*, 2009, 2011; Grealy *et al.*, 2017a).

As we are currently experiencing the Earth's sixth mass extinction, and as the number of species classed as threatened continues to rise (Barnosky *et al.*, 2011; Ceballos *et al.*, 2015; Ceballos *et al.*, 2017), it is an ideal time to use aDNA to learn what we can from extinct species. Seabirds, for example, are more threatened than any comparable groups of birds, with a third of seabird species at risk of extinction, and one half are in, or likely to be in, decline (Croxall *et al.*, 2012). Threats to seabirds include climate change, habitat loss, introduced species, pollution, fishery related impacts (e.g. overfishing of prey, bycatch and entanglement in nets), and also direct exploitation of harvesting eggs, chicks and adult birds (Croxall *et al.*, 2012; Paleczny *et al.*, 2015). Since seabirds play a globally important role in ecosystems (Paleczny *et al.*, 2015) understanding more about the cause of seabird extinction is of high priority. Thus, investigating the population genetics of a recently extinct seabird may yield insights relevant to conservation of contemporary species. The great auk is an exemplar of such a species and this thesis has focused on the evolution and extinction of the great auk, to not only reveal greater insight into an iconic extinct species, but also provide us with information that may benefit extant species.

1.2 The great auk, *Pinguinus impennis*

The great auk (Fig. 1.1), *Pinguinus impennis* Bonnaterre (1790) (traditionally *Alca impennis*, Linnaeus, 1758), is one of very few bird species that has gone extinct in recent historical times (Fuller, 1999), with the last pair ever reliably seen killed in June 1844 (Newton, 1861). During its existence no formally trained naturalist ever studied a living wild bird, leaving many unanswered questions concerning aspects of its life history, evolution and extinction (Olson *et al.*, 1979; Bengtson, 1984).

Figure 1.1 A specimen of the great auk with summer plumage, from the collections at Royal Belgian Institute of Natural Sciences (RBINS 5355) (Credit: Thierry Hubin).

It is well documented that the great auk was hunted, but what impact climate-driven environmental change may have played in its extinction is unknown (Bengtson, 1984; Fuller, 1999). Seabirds are known to be both directly and indirectly impacted by environmental change. For example, changes to sea temperatures can impact prey quality, abundance and distribution, which ultimately impacts seabird survivorship and breeding success (Tuck, 1961; Bengtson, 1984). The archaeological record and written accounts from the time both give detailed evidence for a step-wise pattern of decline in great auk numbers, leading to their eventual extinction. Such conditions provide a valuable and exciting opportunity to study the process of extinction and investigate the effects of environmental change and human induced pressure in one of the few bird species that has gone extinct in recent historical times. The cause of extinction has been investigated in several iconic megafaunal species whose extinction was similarly thought to have been caused by humans. However, the results of such studies found that environmental change was a strong component of their extinction (e.g. woolly rhinoceros (Lorenzen *et al.*, 2011), cave bear (Stiller *et al.*, 2010), Patagonian megafauna (Metcalf *et al.*, 2016)) . This may therefore be the case with the great auk. Using a paleogenetic approach, which allows us to examine changes in population size, population structure and levels of genetic diversity, we aim to investigate the evolution and extinction of the great auk and by gaining a better understanding of the process, it may be possible to better evaluate the threats to extant seabird species arising from the impact of humans and environmental change.

1.2.1 Alternative names and taxonomy

Both during its existence and since its demise, the great auk has been known by many names. Its wide distribution and presence around the shores of numerous counties led to several names for the same bird (Grieve, 1885). Therefore, obtaining knowledge of these various names is important when examining the literature. Many of the names given to the great auk are in reference to its appearance (Grieve, 1885). For example, *Esarokitsok*, was used by the Inuit of Greenland, and translates to meaning stump or small wings. *Brillefugl* and *Brillenalk*, both translate to mean spectacle bird/auk, in reference to the white patches on their heads. Other names refer to its large, hooked, or spear like beak, such as *Geyr-fugl, Geirfugl*, and countless other variations of this word, including *Garefowl*, the name used around Scotland. And, of course, the most common name these days, great auk, is also in relation to its appearance, due to its size in comparison with related species, as is the German term *Riesenalk*, translating to Giant Auk. Perhaps the most interesting name that once belonged to the great auk is *Penguin*. The great auk is said to be the *Original Penguin* or the *Northern Penguin*, the name thought to originate from the Welsh terms for head, *Pen*, and white, *Gwyn*, referring again to the white patches on their heads. It also could come from the Latin, *pinguis*, meaning fat. Either way, the great auk was originally given this name and as sailors travelled south and saw the birds we know today as penguins (members of the *Sphensicidae* family), which had a similar appearance and behaviour, and as the great auk became increasingly less common, the name transferred to them. Other more unusual names assumed to belong to the great auk are *Apponath*, *Apponatz, Apponar* and *Wobble* (Grieve, 1885; Fuller, 1999).

In *Systema natura*, (the renowned work by Linnaeus 1758), the great auk was titled *Alca impennis*, literally meaning the flightless or wingless auk. Since then there were numerous suggestions and changes to the great auk nomenclature, such as *Platus impennis*, based on its relationship with the razorbill and its superficial relationship to penguins. However, in 1856 Banaporte proposed a name that had originally been used by Bonnaterra in 1791, *Pinguinus*. This name was described as 'apt' due to the great auk being the 'original penguin' and so the great auk became *Pinguinus impennis* (Gaskell, 2000).

In terms of its modern scientific classification, the great auk can be found, along with the other twenty-two extant species of auk and murre, in the order Charadriiformes and family Alcidae (Fig. 1.2). Its sister relationship with *Alca*, the razorbill, has long been proposed (Gaskell,

2000; Moum *et al.*, 2002; Pereira & Baker, 2008) but was confirmed using genetic data in 2002 (Moum *et al.*, 2002).

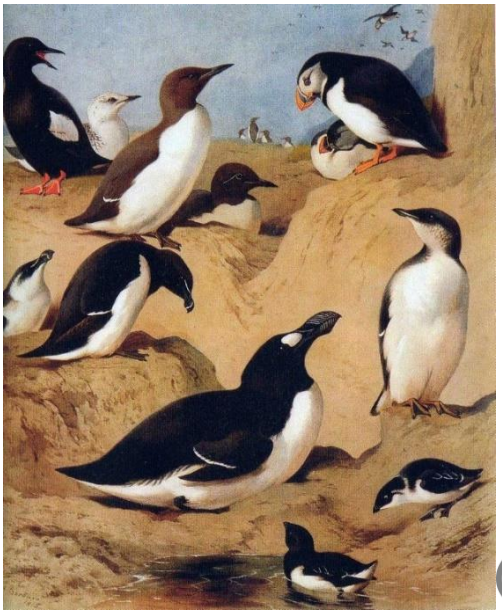

Figure 1.2 A painting of the great auk (centre) with some of its extant relatives from the Alcidae family by Archibald Thorburn (clockwise from top right: Atlantic puffin, thick-billed murre, little auk, great auk, razorbill, common murre, black guillemot). (Image is free from copyright).

1.2.2 Appearance and morphology

The great auk, was a flightless, black and white bird, with a large beak, upright stance and great size, in comparison to other members of the Alcidae family (Fuller, 1999). Its summer appearance is well known, as this was when they were ashore to breed, and most of the specimens in exhibits today were killed during the summer months (Fuller, 1999). Their breeding/summer plumage was black on the upper parts and head, with the exception of the white patch in front of the eye. The upper throat and sides of neck are again black but with a brownish tinge. The feathers of the wings are black but the secondary feathers are white. The under parts are white but the flanks were commonly observed as being grey below the wing (Fuller, 1999). Spending the winter months at sea, descriptions of its wintertime appearance are rare. It is said that the changes from summer to winter plumage were similar to that of close relatives like the razorbill with the brownish- black on the throat and fore neck turning white (Gaskell, 2000). In winter/non-breeding plumage, and in immature birds, the white patch in front of the eye was reduced to a grey line (Montevecchi & Kirk, 1996).

The great auk is described as having a *'full and compact form'* with a short, thick neck and a large head (Gaskell, 2000). The wings were very small, on average 16cm, but despite not being able to support flight were still wing-like in structure. It had webbed feet, that were set far back and toes with small claws (Fuller, 1999; Gaskell, 2000). The bill was large, black and laterally compressed with between 6-12 longitudinal grooves (Montevecchi & Kirk, 1996).

The great auk stood at around 75-80cm (Fuller, 1999), and weighed an estimated 4750g (Smith, 2015)-5000g (Bédard, 1969; Livezey, 1988) making it the largest Alcidae species and the second largest species of Charadriiform, after the extinct *Miomancalla howardae* (Smith, 2015). There is no evidence for sexual dimorphism in plumage (Fuller, 1999), though, it may have occurred in some skeletal elements such as bill and femur (Livezey, 1988). There is, however, evidence of oceanographic-related size differences, with great auks of the North-West Atlantic (i.e. Low Arctic oceanographic zone) being larger than those of the North-East Atlantic (i.e. Boreal oceanographic zone). The size difference was attributed to the trend of an increase in body size with a decrease in sea surface temperatures or increasing latitude, observed in other alcid species (Burness & Montevecchi, 1992) and perhaps an indication of population structure between colonies, as authors suggested gene flow between colonies on either side of the Atlantic was unlikely (Burness & Montevecchi, 1992).

1.2.3 Distribution

The great auk was distributed across the North Atlantic (Fig. 1.3) with a similar distribution to that of the northern gannet (*Morus bassanus*) (Serjeantson, 2001) and not a bird of the Arctic as stated erroneously in early published accounts (Bengtson, 1984; Fuller, 1999). In terms of its distribution, much of where the bird went during the winter months can only be speculated, however, what does exist are records of breeding sites, and it has a relatively rich fossil record. However, it is likely that there were many more colonies than we know of (Montevecchi & Kirk, 1996), and equally, inferring range based on the fossil record holds some risk as it does not mean great auks lived or bred in a particular locality, as individuals may have been killed elsewhere and traded (Fuller, 1999; Serjeantson, 2001).

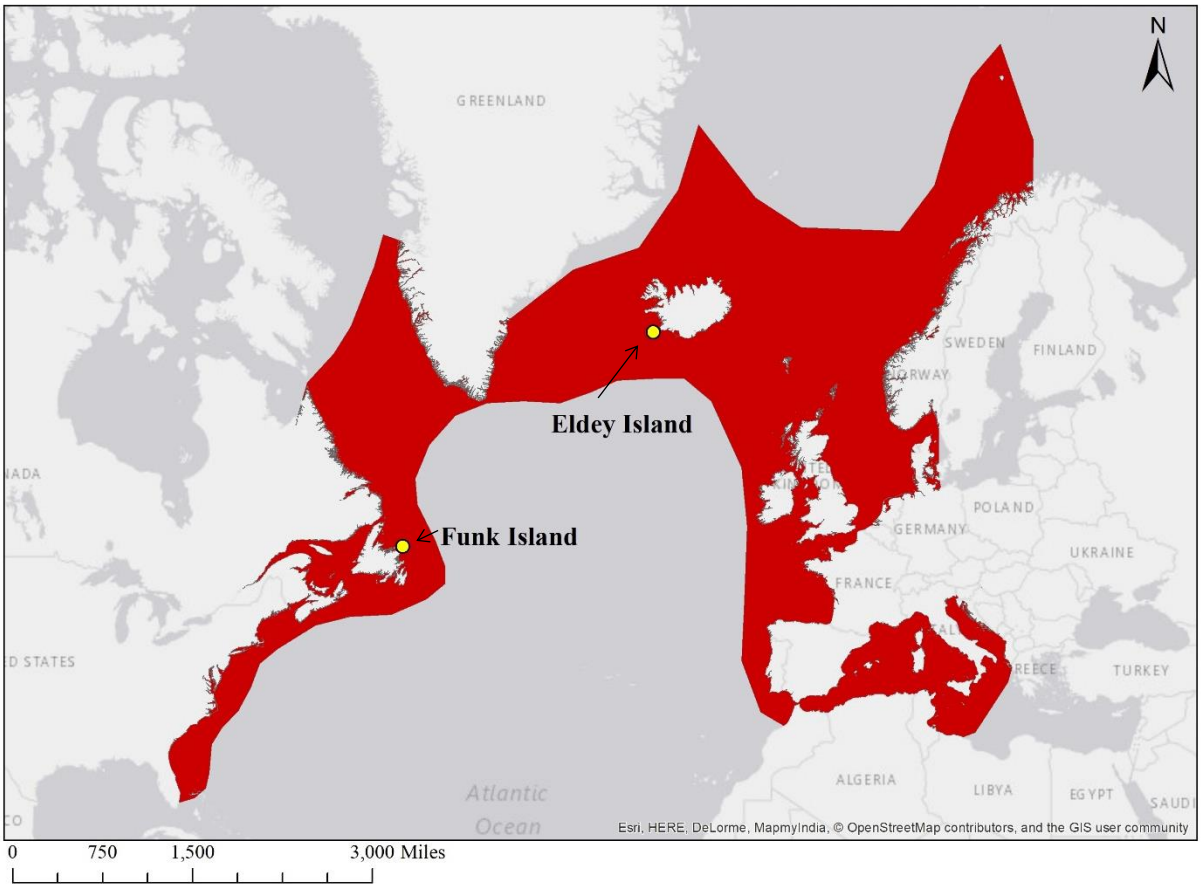

Figure 1.3 The former distribution of the great auk, as defined by BirdLife International/IUCN (2016), however, this study includes samples from sites beyond the distribution boundary in Norway. Sites indicated by yellow dots represent key great auk locations: Funk Island was the site of the largest great auk breeding colony, Eldey Island is the location where the last two individuals ever reliably seen were killed in June 1844.

In the Western Atlantic, breeding colonies could be found along the east coast of North America. Many islands around Newfoundland are synonymous with the great auk, with at least 11 locations around Newfoundland and Labrador referencing the name penguin. Funk Island, which lies ~60km from Newfoundland, was home to the largest colony of great auks (Montevecchi & Kirk, 1996). It was also found on islands in the Bay of St. Lawrence, Cape Breton and Cape Cod (Massachusetts). There is archaeological evidence of great auks as far south as Florida on the western coast of the Atlantic, although not in the numbers found around Funk Island (Hay, 1902; Brodkorb, 1960; Bengtson, 1984; Serjeantson, 2001). Great auks also bred in small numbers in Greenland. Wintering birds could also be found off Greenland's coasts, which may have come from the breeding colonies of Iceland or Funk Island, depending on directions of migrations (Meldgaard, 1988). Breeding colonies could be found on many of Iceland's isolated skerries (Bengtson, 1984), the two most famous being Geirfuglasker and

Eldey Island. It also bred on the islands off Scotland, including St Kilda in the Outer Hebrides and the Orkneys and possibly the Isle of Man between Wales and Ireland. There are also several finds of great auk bones in sites around the British Isles, many of which show evidence of butchery (Serjeantson, 2001). Great auk bones have been found in numerous sites throughout Scandinavia, including kitchen middens in Denmark (Grieve, 1885; Meldgaard, 1988) and several sites of various ages in Norway (Hufthammer, 1982; Bengtson, 1984). Moving southward through Western Europe, there is evidence at several sites, from various geological periods including southern Italy, Portugal, Spain, France (Upper Pleistocene and Holocene) and even Morocco (Holocene) (Serjeantson, 2001; Campmas *et al.*, 2010). The great auks' distribution undoubtedly changed throughout time, possibly in line with climatic changes, with both northward and southward movements evident. Their decline from many locations is also apparent by the presence and absence in various dated levels of archaeological sites (Bengtson, 1984; Serjeantson, 2001).

1.2.4 Life History: Behaviour and breeding

The great auk was a mysterious bird, with many details of its life history unknown, particularly what it did during its ten months of the year spent at sea (Bengtson, 1984; Fuller, 1999). The great auk was an expert swimmer and diver, with its wings superbly adapted to subaqueous flight, although they were still true wings and not flipper like as with penguins (Bengtson, 1984). Even its large size and mass would have been advantageous as it reduced buoyancy and allowed for greater diving depths to be reached (Livezey, 1988). The bill is well designed for hunting fish, with features for catching, carrying and slicing the fish identifiable (Fuller, 1999). Otto Fabricius (1780) performed dissection of adult and chick great auks in order to study their diets. However, it has been expressed by some of the most accredited great auk researchers that Fabricius had made an error in identification and therefore perhaps his work should be treated with caution. A dissection of a chick showed stomach contents contained shore plants such as *Sedum rhodiola*, while in the stomachs of adults, Fabricius found shorthorn sculpin (*Cottus scorpius*) and lumpsucker (*Cyclopterus lumpus*) (Fuller, 1999). Reports of great auks held captive, such as the young bird captured in Waterford Harbour, Ireland, are recorded as being fed sprats (*Sprattus sprattus*) and herrings (*Clupea harengus*) (Fuller, 1999). More recent research into feeding ecology has employed techniques which analyse bones of great auk (Hobson & Montevecchi, 1991) or look in detail at soil where the auks lived (Olson *et al.*, 1979). Olson et al. (1979) used the soil that remained on bone samples, initially collected

by F. A. Lucas from Funk Island on the Grampus expedition of 1887, to infer the prey of great auks. By examining the soil for remains of prey items they identified several fish species including shad, capelin, flatfish and three-spined-stickleback. From this they inferred that the auks fed in waters of up to 18m depth and within 2km of the shore; however, even authors stated that the conclusions were drawn from *"very slender evidence"* (Olson *et al.*, 1979). Hobson & Montevecchi (1991) used isotopic analysis of bone collagen of great auks from Funk Island to determine trophic relationships and found similar results to Olson et al. (1979). Hobson & Montevecchi (1991) suggested that prey was eaten from several trophic levels, ranging from trophic level 3 i.e. crustaceans, to trophic level 5, piscivorous and benthic fish. The results were further interpreted to suggest that the lower trophic level results were from young birds feeding on krill (*Euphausiacea* species), while the higher 15N values were from the older birds who fed on piscivorous and benthic fish (Hobson & Montevecchi, 1991).

Despite many of the accounts of the great auks being recorded during their breeding season, there are actually very limited details regarding their breeding habits (Bengtson, 1984). One of the most comprehensive reviews of great auk breeding biology is by Bengtson (1984), though he states that due to the limited information available, there is much speculation and many details are influenced by knowledge of living relatives. Being flightless, great auks were somewhat limited in their choice of breeding sites. They required either a suitable landing place from which they could scramble up, or use the surf to get ashore (Bengtson, 1984). Montevecchi & Kirk (1996) describe three factors that determined suitability of breeding sites: 1) distance from mainland/humans/predators (e.g. polar bears), 2) physical topography, 3) proximity to feeding grounds. However, they also comment that low-lying islands may not have been a necessity, as other flightless species, such as rock hopper penguins, can climb cliffs (Montevecchi & Kirk, 1996). In terms of timing of breeding, there are various dates and timings suggested of when they first came ashore and how long they spent there. The dates also seem to vary by location. For example, approximate dates for when great auks were ashore in St Kilda, Scotland, is 12 May-26 June. However, in Iceland it was reported that eggs were collected and birds killed in late July/early August. It is thought that only a single egg was laid, and not replaced if broken or lost. The egg was much larger than other Alcids (great auk volume 3000cm^3, razorbill ~81-83cm^3, common murre 96cm^3) and described as being *'all yolk'* (Bengtson, 1984). Both sexes had a single brood patch, suggesting that incubation was by both males and females. Incubation is estimated to have taken around 40-45 days. A short fledging period of 9 days has been proposed based on other published information about how

long great auks were on land and their incubation time (Bengtson, 1984). Bengtson (1984) argues that a short fledgling period was not unlikely, as it could explain why there are few descriptions of chicks and none in museum collections. He also suggested that due to the fact great auks were clumsy on land and the effort it took to forage for themselves and a chick, and take that food back to them onshore, it would have been beneficial to get the chick to sea quickly so it could feed it more easily (Bengtson, 1984). Houston et al. (2010), suggest that the great auk followed an intermediate strategy for chick rearing, i.e. young were raised to around 20% adult mass, as does the razorbill (Houston *et al.*, 2010). As there is no record of breeding age, estimates based on related species suggest that they did not breed until ~4-7 years old (Bengtson, 1984). It is likely that as with extant auk species, reproduction would be unsuccessful in years of food shortage, together with sea ice that likely also impacted breeding success (Montevecchi & Kirk, 1996). It is thought that the great auk was a long lived species, based on relationship between body mass and survival for other Alcids (Birkhead, 1993), with a life span of 20-25 years (Bengtson, 1984). Finally, Bengtson (1984) suggests it is possible that the great auk did not breed every year, and similar to other closely related species such as the razorbill, showed a high degree of nest site and mate fidelity, and so are presumed monogamous (Bengtson, 1984).

The great auk is known to have been a very social species. This is evident from recorded behaviour and observations such as head bobbing and shaking, its white eye patch, white bill markings, and the yellow skin on the inside of their mouth, all likely to have been used in social displays as in other alcids (Bengtson, 1984; Fuller, 1999). It is also evident from the large breeding colonies that formed each summer on island such as Funk Island, and the large rafts of birds observed at overwintering sites (Fuller, 1999). Many seabird species live in colonies as there are a number of advantages of living in large groups. For example, increase in foraging efficiency, increased chances of finding a mate due to more balanced sex ratios, and anti-predator defence (Schippers *et al.*, 2011). It is unknown what species did prey on the great auk, however, it is possible that its predators included polar bears, killer whales and white-tailed eagles, that could have preyed on adult birds (as they had a considerable amount of subcutaneous fat) or eggs (Montevecchi & Kirk, 1996). Living in large colonies would have caused a 'diluting effect' which reduces the risk of an individual chick/egg being predated, and 'anti-predator defence' due to more individuals to act as a predator alert system (Schippers *et al.*, 2011).

1.2.5 Genetic research

To date, genetic research on the great auk is limited. DNA sequences have been used to resolve its taxonomic position within the Alcidae, and to calculate dates for its divergence from the razorbill. In 2002, Moum et al. sequenced the first great auk DNA from a ~180-year-old mounted specimen. A 4.2-kb region of mitochondrial DNA was amplified using Polymerase Chain Reaction (PCR) amplification. The sequence was used to resolve the phylogeny in the Atlantic Alcidae and confirm the sister relationship between *Pinguinus* and *Alca torda*, the razorbill (Fig. 1.4). They also suggested that the great auk originated late in relation to the other divergences (Moum *et al.*, 2002). In 2016, the complete mitochondrial genome of the great auk was published online (GenBank: KU158188). It was sequenced from a mounted specimen collected in 1831, from Eldey Island, Iceland, currently held in the Natural History Museum, Oslo (Anmarkrud & Lifjeld, 2017).

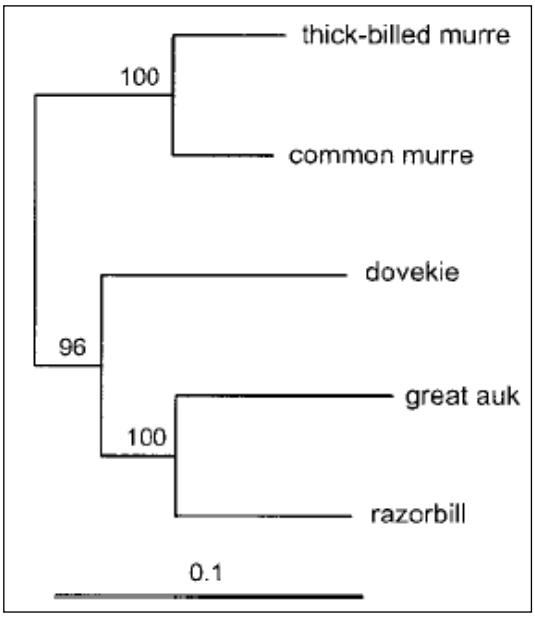

Figure 1.4 Phylogenetic tree of Atlantic alcids from Moum et al. (2002) showing the relationship between the great auk and razorbill. The phylogenetic tree was created using maximum likelihood analysis of 3,140 nucleotide sites of mitochondrial DNA. The black guillemot was used for outgroup rooting and node values represent bootstrap replication scores based on 100 resamplings (Moum *et al.*, 2002).

Estimates of divergence time between great auk and the razorbill have yielded varying results. Using genes sequenced by Moum et al. (2002), Baker et al. (2007) calculated the divergence time to be 22.5million years ago (± SD 4.4, 95% Credible Interval 14.9-31.3) (Baker *et al.*,

2007). In 2008, Pereira & Baker published another set of dates, this time calculating divergence time to be 18.7mya, 24.2mya, 27.9mya, depending on the method used (Pereira & Baker, 2008). In the same study which looked at 21 of the 23 extant auks, they suggested that the most recent common ancestor of *Alca*, *Alle*, *Pinguinus* and *Uria*, invaded the Atlantic Ocean during the Eocene/Oligocene, between 40-30mya (Pereira & Baker, 2008). Smith & Clarke (2015), used a combined approach of phylogenetic and morphological data to investigate the systematics and evolution of the Pan Alcidae (true auks, dovekies, murres, murrelets, guillemots, auklets, puffins and the extinct mancalline/Lucas auks). The study calculated divergence of the great auk and razorbill to be considerably younger than that previously suggested (Baker *et al.*, 2007; Pereira and Baker, 2008), at mean age of 11.7mya (95% HPD 7.8-15.2mya) (Smith & Clarke, 2015). The study also suggested the Pan-Alcids diversified during periods of relative warmth (e.g. Middle Miocene Climatic Optimum 16-11mya) and underwent extinction events at times of cooling (e.g. Pliocene Pleistocene Climatic Transition 3-2mya). The authors do, however, note that such speculation seems counterintuitive for alcids, who are reliant on cold water upwelling (Smith & Clarke, 2015).

An additional, and very different, interest in great auk genetics has also recently arisen as the great auk has been proposed as a candidate for de-extinction (Brand, 2016). De-extinction projects aim to literally regenerate extinct species (Richmond *et al.*, 2016), and, although such propositions are controversial and will face numerous challenges, debate and ethical concerns, they are gaining interest from both certain segments of the scientific community and general public (Richmond *et al.*, 2016; Grealy *et al.*, 2017a). However, bird reproductive physiology makes genetic engineering much more difficult than with mammals as the oocytes are fertilised shortly after release and then covered in a hard shell (Richmond *et al.*, 2016). It has been reported that there are several known issues with IVF and the zygote transfer process which may make cloning steps unachievable (Richmond *et al.*, 2016; see Sang, 2004 for more information). Additionally, genome editing using CRISPR/Cas9 would be more difficult or perhaps not possible as this requires manipulations of the cell line, which need to be distinguished, and preservation of the cell for the number of years required to make the changes may be problematic (Richmond *et al.*, 2016). Nevertheless, such issues have not precluded meetings to discuss the possibilities of bringing the great auk back from extinction (Brand, 2016), and also led to the recent sequencing of a full draft great auk genome (Niemann, Gilbert, *et al*., unpublished).

1.2.6 The demise of the great auk

The demise of the great auk is well known; this is in part due to the role that humans played in its extinction. The archaeological record and numerous written accounts that exist regarding great auk hunting are testimony to the scale at which the killings of great auks occurred, especially on Funk Island. Being flightless and adapted to a life in the water, the great auk was clumsy and awkward on land (Grieve, 1885; Fuller, 1999). This made them especially vulnerable and an easy target for hunters when they came ashore each summer to breed. In prehistoric times, they were hunted by Beothuk Indians in North America (Fuller, 1999; Gaskell, 2000), Inuit's of Greenland (Meldgaard, 1988), Icelanders and Scandinavians (Bengtson, 1984), during various periods in the British Isle (Serjeantson, 2001) and possibly even Neanderthals (Halliday, 1978). However, from ~1500CE, when the Europeans discovered the rich fishing grounds of Newfoundland, year on year fleets of 300-400 ships could be found there (Steenstrup, 1855; Grieve, 1885; Bengtson, 1984). The fishermen showed a strong preference for choosing fishing stations near to seabird colonies, as they not only used them as navigational guides and as indication of fish stocks, but also as an additional food source (Pope, 2009).

The birds of Funk Island, and other islands around Newfoundland coast, were exploited in many ways and the immense level at which the hunting occurred is evident from reports such as the following from Jacques Cartier, 1534 quoted in Gaskell (2000) *'We put into our boats so many of them as we pleased, for in less than one hour we might have filled thirty such boats of them'*(*page* 57), *'in less than half an hour we filled two boats full of them as if they had been with stones: so that besides them that we did eat fresh, every ship did powder and salt four or five barrels full of them'* (*page 52)* (Gaskell, 2000). Such high level of capture is reiterated by several others, for example a Franciscan friar, Gabriel Sagard Theodat (1624) (quoted in Gaskell, 2000 page 57) *'…certain kinds there which cannot fly at all, and which one can easily overpower by blows with club, as did the sailors of another ship, who before us had used their shallop* [sic] *for that purpose, and several barrelfuls* [sic] *of eggs, which they found in the nests'* (Gaskell, 2000). Richard Whitbourne (1622) described how hundreds of great auks at a time were herded down gangplanks into boats (Fuller, 1999). In addition to those that were killed on land, there are reports of seabird species, including guillemots and petrels, in addition to the great auk being *'fished'* with a line while at sea (Pope, 2009). It may also be likely that great auks could have been caught in nets that would have frequented the waters of the fishing

grounds and drowned, as is seen in many extant diving birds (Piatt & Nettleship, 1985, 1987; Montevecchi & Kirk, 1996). As well as harvesting eggs and adult birds for food, great numbers of the young were used as bait (Grieve, 1885). Pope (2009) estimates that 50,000-60,000 adult seabirds (not just great auks) were taken for the use as bait, from every exploitation zone, each year. In addition to the eggs that were collected, he estimates 20,000 adult birds were killed, if each fisherman was eating half a seabird once a week for eight weeks (Pope, 2009). Several writers (Kirkham & Montevecchi, 1982; Gaskell, 2000; Pope, 2009) suggest that it was not hunting for food source that caused the decline of the great auk around Newfoundland waters. They propose instead that it was the development of the commercial hunt for feather trade that eliminated the population (Kirkham & Montevecchi, 1982; Gaskell, 2000; Pope, 2009).

Throughout their distribution, local population extinctions can be traced with evidence of fewer individuals in the various dated layers (Serjeantson, 2001). The decline occurring in an almost step-wise fashion with huge numbers initially being harvested to only few individuals as numbers dwindled. In many places, dates of last recorded sightings of individuals and even numbers of individuals killed are available (Fig. 1.5) (Grieve, 1885; Fuller, 1999; Gaskell, 2000).

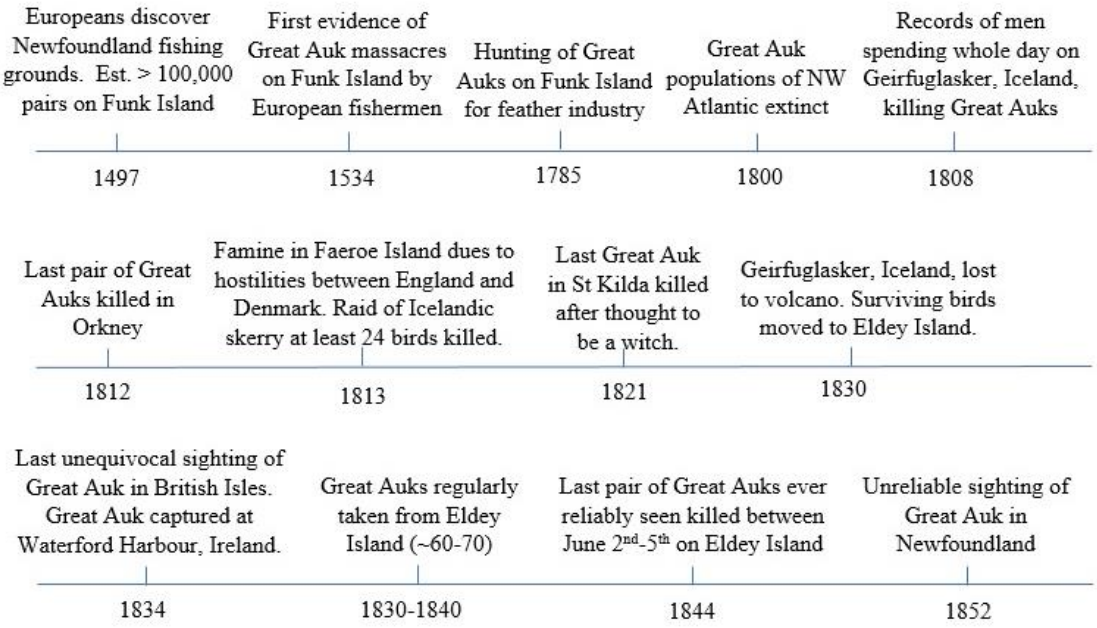

Figure 1.5 A timeline of key events that took place in the last few hundred years of the great auk existence, from the discovery of the fishing grounds of Newfoundland to the killing of the last pair ever reliably seen in June 1844 on Eldey Island, Iceland (Grieve, 1885; Fuller, 1999; Gaskell, 2000).

With populations of the North-West Atlantic decimated by around 1800 (Bengtson, 1984; Fuller, 1999), it was on the skerries off the south-west coast of Iceland that events leading to the end of their existence played out. The population numbers here are thought to have also been low by this time, with perhaps only a few hundred birds visiting these islands annually (Bengtson, 1984). The birds on these Icelandic skerries were a food source to the Icelanders for centuries and raids of island for eggs and meat are known to have occurred during the years of good weather when they could reach and land on the islands (Newton, 1861; Fuller, 1999).

In 1830, one of the great auks last strongholds, Geirfuglasker, was lost due to volcanic activity, and so many of the surviving birds had moved to nearby Eldey Island. This island was closer to the main land, increasing the risk of great auks to hunting by humans (Fuller, 1999). By this time, the great auks were also sought after to be put on display in exhibitions (Fuller, 1999), although they would have undoubtedly been killed for food regardless of the exhibition trade. In 1830, Brandur Guðomundsson led two voyages to Eldey. Twelve birds were killed during the first and eight during the second. At least twelve of these birds were sold to merchants (Newton, 1861). In 1831, twenty-four were captured, and all skinned by the same woman, Sigrida Thorlaksdotter. In 1833, 13 birds were taken and in 1834, 9 birds and 8 eggs were obtained. In 1840/1841, three skins and eggs were taken from Eldey (Newton, 1861). Following a three-year period of no recorded captures of great auks, Carl Siemsen commissioned an expedition to Eldey to search for any remaining birds. Between 2nd and 5th June 1844 the expedition reached Eldey Island where two great auks were observed amongst smaller birds inhabiting the island. Both auks were killed and their broken egg discarded. The birds, though, were never to reach Siemsen. The expedition leader sold them to Christian Hansen, who then sold them to the apothecary Möller, in Reykjavik, Iceland. Möller skinned the birds and sent them, as well as their preserved body parts, to Denmark (Newton, 1861). These were the last two great auks ever reliably seen and the organs of these birds now reside in the Natural History Museum of Denmark. The location of the skins from those individuals remained a mystery until genetic research employed as part of this thesis was able to partially resolve the mystery as a match between the organs of the male individual and the skin housed in the RBINS in Brussels was found. The location of the skin from the female bird remains unsolved but speculations as to its location have been made (Thomas *et al.*, 2017) (Chapter 5).

Records of individuals observed after 1844 have been discussed in the literature (e.g. 1848 Vardø (Newton, 1861; Fuller, 1999), 1852 Newfoundland (Newton, 1861; Grieve, 1885;

Fuller, 1999) and BirdLife International/ IUCN recognises the last sighting as 1852 (BirdLife International, 2016b). However, there is some degree of uncertainty about these sightings (Grieve, 1885; Fuller, 1999) and in general June 1844 is the date quoted as the extinction of the great auk.

As well as these documented declines in abundance of the great auks, some authors have eluded to the fact that climatic events may have also played a part in its extinction (Bengtson, 1984). Tuck, (1961) commented on the effect climate change had on alcid populations stating that there was good evidence that it could both directly and indirectly impact this family (Bengtson, 1984). Alcid populations are dependent on the abundance and availability of food. If their prey is affected by sea water temperatures, resulting in changes in distribution or abundance, then this impacts the alcids as the lack of food can reduce breeding success (Bengtson, 1984). To quote Bengtson, (1984) (page 10), '*In the absence of more detailed information about rate of decline of the bird populations, hunting pressure and environmental changes, we cannot separate the effects of hunting and that of climate change'.* He also commented that with the bird being flightless it would have required safe breeding areas with abundant food sources in order to successfully breed, he noted that *'traits in life history probably made it more vulnerable to climate change*' (Bengtson, 1984) (page 10).

1.2.7 Why study the great auk?

While the review of literature shows that there has been much written about the great auk, it also highlights that there are gaps to the knowledge and reveals many unanswered questions regarding its evolution, existence and extinction. For example, was the great auk at risk of extinction and in decline prior to the period of intense hunting that began in ~1500CE? Or, did climatic events such as the 'Little Ice Age' contribute towards their extinction or are humans alone responsible? These questions, plus many more, can start to be addressed using ancient DNA.

1.3 Ancient DNA

Ancient DNA has been broadly defined as *'the retrieval of DNA sequences from museum specimens, archaeological finds, fossil remains and other unusual sources of DNA'* (Pääbo *et al.*, 2004) (page 645). It allows us go back in time and explore features such as past trends in

population connectivity, size and distribution, levels of genetic diversity and response to environmental change, hunting or species introductions (Leonard, 2008; de Bruyn *et al.*, 2009, 2011; Grealy *et al.*, 2017a). Over the past three decades, the field of aDNA has grown considerably, from sequencing a small section of mitochondrial DNA (mtDNA) from the Quagga, an extinct form of the plains zebra (Higuchi *et al.*, 1984), to whole genome sequencing from samples up to 735,000 years old (oldest complete genome sequenced from a horse preserved in permafrost in the Yukon) (Orlando *et al.*, 2013). Since its modest beginnings, aDNA has addressed a diverse range of ecological and evolutionary questions, providing insight into countless species' pasts, including our own. However, it is an area of research that has faced many challenges and pitfalls, but nevertheless, advances in technology have allowed the field to continue to grow. This section, while discussing the broader aspects of aDNA, provides a more detailed insight into the use of ancient DNA in avian studies, and how we apply a paleogenetic approach to investigate the evolution and extinction of the great auk.

1.3.1 Sources of aDNA

Sources of aDNA are typically those that have not been preserved for use in genetic research (de Bruyn *et al.*, 2011). Following the death of an organism, DNA molecules begin to degrade rapidly, become damaged and chemically altered (Pääbo *et al.*, 2004). Therefore, samples of aDNA characteristically contain short fragments of chemically altered DNA, with a low amount of endogenous (target) content, often with high levels of contamination from exogenous DNA (Pääbo *et al.*, 2004; Willerslev & Cooper, 2005; Rizzi *et al.*, 2012; Green and Speller, 2017). However, if samples have become frozen quickly (such as those in permafrost) or desiccated rapidly, then the processes that cause DNA damage and alteration are inhibited, leading to better preservation (Hofreiter *et al.*, 2001; Dabney *et al.*, 2013). Numerous environmental conditions affect DNA preservation, such as temperature, salinity, pH and humidity (Dabney *et al.*, 2013). To date, the vast majority of studies have utilised samples from environments that promote good DNA preservation such as, those found in permafrost, (horse (Orlando *et al.*, 2013)), caves (cave bear (Dabney *et al.*, 2013), Denisovan (Meyer *et al.*, 2012)) and cold environments (Antarctic penguins (Lambert *et al.*, 2002)). However, advances in methods and technology are not only facilitating a wider range of novel sources of DNA to be utilised, but also those from environments not typically thought to be good for DNA preservation, such as warm tropical and subtropical regions (camels (Mohandesan *et al.*, 2017), Bahamas giant tortoise (Kehlmaier *et al.*, 2017), Caribbean rice rats (Brace *et al.*, 2015),

Yucatan rats (Gutiérrez-García *et al.*, 2014), and even enslaved African humans from the Caribbean (Schroeder *et al.*, 2015)).

The first decade of aDNA research primarily used soft tissue from museum collections(Higuchi *et al.*, 1984; Pääbo, 1985; Pääbo *et al.*, 1988), but by the 1990s extraction from bone was shown to be successful and this expanded opportunities in aDNA analysis (Hagelberg *et al.*, 1989; Horai *et al.*, 1989; Green & Speller, 2017). Studies using bones and teeth dominate the ancient DNA literature, however, there are an increasing number of studies being published that utilise novel sources (reviewed in (Green & Speller, 2017). For example, herbarium collections and archaeobotanical remains, cultural archives or stone and ceramics, collagenous or keratinous tissues, paleofaeces and coprolites (Green & Speller, 2017). One important point to make is that while there may be numerous sources of DNA, much of this material is available in small amounts, rare, and invaluable, especially as many aDNA studies focus on extinct species (Payne & Sorenson, 2003). In the early days the rarity and importance of samples was a problem, as huge amounts of sample were needed for protocols (Der Sarkissian *et al.*, 2015), and methods of sampling were classed as destructive (Payne & Sorenson, 2003). However, as protocols and technology has advanced, less starting material is required, and non-destructive methods of sampling have developed (Rohland *et al.*, 2004; Gilbert *et al.*, 2007; Shepherd, 2017; Teasdale *et al.*, 2017). Additionally, museums acceptance of the value and importance aDNA studies can have in adding to specimen and scientific knowledge, has made many more aDNA studies feasible (Payne & Sorenson, 2003).

Within avian aDNA studies, Grealy et al. (2017a) revealed that in 156 studies that extracted avian aDNA, 48.1% used museum skins, 50% from single-source bone and 13.5% from egg shell, coprolites, feathers, sediment or bulk bone (NB. This equates to over 100% as some studies used more than one sample type) (Grealy *et al.*, 2017a). With regards to the great auk, sources of aDNA range from bones collected from archaeological excavations (e.g. Greenland (Meldgaard, 1988), Funk Island (Lucas, 1890), British Isles (Serjeantson, 2001), Scotland (Grieve, 1885), Spain (Espolosin Elorza, 2014), Norway (Hufthammer, 1982) , the Netherlands (Groot, 2005)), 78 mounted skins and 75 eggs in museum collections and the organs from the last individuals stored in spirits (Fuller, 1999). For the purpose of the present population genetic study, it is important that sample provenance is known. For much of the bone material this is the case, with the location data available from archaeological sites, and also in many cases from stratigraphic information on time period and estimated age. However, only 20 of

the mounted skins have a known country of origin, for the others the location information is listed as 'unknown' or 'unknown but probably…' (Fuller, 1999), therefore, we shall only sample skins if information is known. Whilst avian egg shells have shown to be a good source of aDNA for many species (Oskam *et al.*, 2010; Grealy *et al.*, 2017b; Jain *et al.*, 2017), for the purpose of our study, eggs will not be used due to their relative rarity and importance of these specimens. Indeed, there are other sources of material that will provide an extensive dataset.

1.3.2 The challenges

Since the beginning, aDNA analysis has faced several challenges due to many of its defining characteristics. Contamination by exogenous DNA is one of the biggest issues and this was demonstrated in many of the early studies. For example, mtDNA was said to be extracted from dinosaur bones (Woodward *et al.*, 1994) and insects in amber (Cano *et al.*, 1993), however, further investigation has led to these studies being dismissed as sequences were shown to be as a result of contamination (Pääbo *et al.*, 2004). In order to ensure the generation of authentic and meaningful results various measures are adopted to minimise significantly such issues. Cooper & Poinar (2000) proposed a list of nine criteria to follow to improve the reliability and authenticity of results. These criteria included, isolation of work areas, using negative controls, independent replication and appropriate molecular characteristics (i.e. length of fragments) (Cooper & Poinar, 2000). However, as discussed in Gilbert et al. (2005) the majority of researchers do not commit to using all nine of the criteria for every study (Gilbert *et al.*, 2005). Gilbert et al. (2005) point out that the criteria were produced to assist in determining the authenticity of a study and the list is adapted to fit each study and situation, as each study faces different problems and issues, depending on the organism and question being investigated. For example, there is greater risk of contamination by related DNA that will jeopardise the reliability of studies of humans, pathogens or microorganisms and domestic species, whereas for well-characterised extinct species, such as moas, there is a reduced risk as the contamination is more easily identifiable (Gilbert *et al.*, 2005).

The issues of poor preservation, short fragments and contamination have been addressed by various developments in methods and technology. Initially, it was the introduction of the Polymerase Chain Reaction (PCR) that changed the field of aDNA, making it possible to produce unlimited copies of DNA (Pääbo *et al.*, 2004), produce large DNA sequences from using overlapping PCR fragments (Hofreiter *et al.*, 2015) and extending the time depth that

could be explored up to several tens of thousands of years (Knapp & Hofreiter, 2010). However, the development of next generation sequencing (NGS) technologies really revolutionised the field (Knapp & Hofreiter, 2010). This technology made it possible to sequence billions of reads simultaneously (Hofreiter *et al.*, 2015). It also allowed for the sequencing of very short molecules, which ultimately increases the number of endogenous ancient molecules as opposed to the long contamination molecules (Knapp *et al.*, 2012). Interestingly, in the recent review of avian aDNA studies, Grealy et al. (2017a) found that over 78% studies used PCR and Sanger sequencing, despite the notable advantages of NGS. However, they comment that there has been a marked increase in recent years of the use of NGS in avian aDNA studies (Grealy *et al.*, 2017a).

Due to the degraded nature of DNA in ancient samples, early studies were limited to short barcoding genes of mitochondrial DNA (Knapp & Hofreiter, 2010; Grealy *et al.*, 2017a). Although NGS has facilitated palaeogenomic studies (Hagelberg *et al.*, 2014; Der Sarkissian *et al.*, 2015), aDNA studies are still dominated by mtDNA and mitochondrial genomes (mitogenomes). This is due to the higher copy numbers in mtDNA compared to nuclear DNA (nuDNA) (typically 1000 times more mtDNA than nuDNA (Rizzi *et al.*, 2012), which is therefore more likely to be amplified (Der Sarkissian *et al.*, 2015; Grealy *et al.*, 2017a). Whilst mtDNA and mitogenomes are frequently used as the loci of choice due to the easier nature of working with than nuclear, an increasing number of studies are aiming to use both mitochondrial and nuclear data to address research questions, as both loci have strengths and weaknesses (Grealy *et al.*, 2017a), and may be complementary. As well as NGS, another technological advancement that has transformed the field of aDNA is hybridisation capture. The method has somewhat replaced the use of PCR, and involves capturing the endogenous DNA within a sample by using bait, and non-targeted contaminant sequences, are washed away (Knapp & Hofreiter, 2010).

The development of material specific extraction protocols has also improved the success of aDNA studies. For example, from bone and teeth (Rohland & Hofreiter, 2007), or keratinous tissues (Campos & Gilbert, 2012), and those which allow for the shortest fragments to be obtained (Dabney *et al.*, 2013). Following extraction, and prior to shotgun sequencing or, capture of regions of interest, followed by NGS, DNA fragments are built into sequencing libraries. Traditionally, aDNA library preparation was performed using methods originally developed for modern DNA and used double-stranded DNA (e.g. 454 Life Sciences method of

ligating two adapters to blunt-end repaired double stranded DNA, or Illumina Y shaped adapter with T-overhang ligated to both end of DNA fragments that were manipulated to carry A overhangs) (Gansauge & Meyer, 2013). However, the development of library preparation methods specifically designed for aDNA take into account the characteristic features such as low endogenous content and short fragments increases effectiveness. A single-stranded library (SSlib) preparation published by Gansauge & Meyer (2013) addressed many of the issues of the double-stranded library (DSlib) build methods (Gansauge & Meyer, 2013). They report that using the SSlib preparation increased the sequence yield of a Denisovan sample by at least six fold, when compared to the 454 method, described above (Gansauge & Meyer, 2013).

In order to ensure the highest levels of reliability in the present study, all laboratory work prior to PCR amplification was carried out in designated aDNA facilities and no-template controls to screen for contamination were used. To increase chances of successfully extracting and sequencing aDNA from the great auk, we employed the latest methods and protocols, most suitable to sample type and those which target the smallest, most degraded fragments typical of aDNA. We employed DNA hybridisation enrichment to increase the chances of sequencing target DNA, a single stranded library preparation method and NGS techniques.

1.3.3 Using aDNA to investigate demographic histories

To date, aDNA studies have encompassed, but are no means limited to, resolving issues of taxonomy and revising phylogenies (e.g. equids (Orlando *et al.*, 2009), ratites (Mitchell *et al.*, 2014), eagles (Bunce *et al.*, 2005), mammoth (Chang *et al.*, 2017)), investigating species domestication (e.g. cats (Ottoni *et al.*, 2017), chickens (Xiang *et al.*, 2014), pigs (Larson *et al.*, 2007)), and revealing insights into the evolution of extinct and extant species, including our own evolutionary past (Fu *et al.*, 2013, 2015; Kuhlwilm *et al.*, 2016). In addition to these uses, ancient DNA has been used to examine the demographic history and cause of extinction, in several species, making it a very appropriate tool in our great auk study.

In 2004, Shapiro et al. reconstructed the genetic history of the steppe bison through the use of BEAST software (Shapiro *et al.*, 2004). This study revolutionised how we explored a species past. In addition to this, Bayesian Skyline Plots, as a method of coalescent-based inferences, have also allowed for population dynamics to be explored (Drummond *et al.*, 2005). Since the publication of these papers, numerous studies have examined species' genetic past and

demographic history to reveal significant insights on patterns and dynamics of extinctions, evolutionary history and the impact of environmental change and/or humans on their existence. However, such studies have mainly focused on iconic megafauna from the Pleistocene. For some species, such as the musk ox (Campos *et al.*, 2010; Lorenzen *et al.*, 2011), woolly rhinoceros (Lorenzen *et al.*, 2011) and collared lemmings (Prost *et al.*, 2010; Brace *et al.*, 2012), climate change has impacted past population dynamics (Fig. 1.6). For others, like the steppe bison (Shapiro *et al.*, 2004; Lorenzen *et al.*, 2011), cave bear (Stiller *et al.*, 2010), wild horse (Lorenzen *et al.*, 2011), Patagonian megafauna (Metcalf *et al.*, 2016), a combination of factors likely contributed to species decline. Studies of the New Zealand moa on the other hand found that prior to human arrival, populations showed genetic stability, with no evidence that moa were in decline suggesting that it was too rapid to have left a detectable trace in their genetics (Fig. 1.7). Therefore, suggesting the cause of extinction in this case is likely to be hunting and habitat loss (Rawlence *et al.*, 2012; Allentoft *et al.*, 2014).

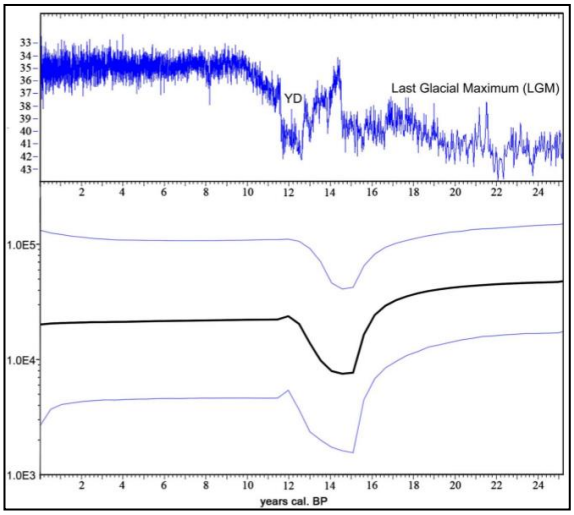

Figure 1.6 Figure from Prost et al. (2010) study investigating the demographic history of the collared lemming (*Dicrostonyx torquatus*). The upper panel displays the climate history as derived from the GISP2 ice-core and the lower panel shows the demographic history of the collared lemming during the last 25,200 yrs. The Bayesian Skyline Plot, used to reconstruct the demographic history, shows the median estimate (thick black line) of the female effective population size (fN_e) (y-axis) over time and the blue lines indicate the 95% highest posterior density intervals (HPD)(Prost *et al.*, 2010). This figure shows a decline in population size during a period of climate change at 14.5kyrs BP. If the great auk had been affected by climatic events such as the 'Little Ice Age' we could expect to see a similar result in our data.

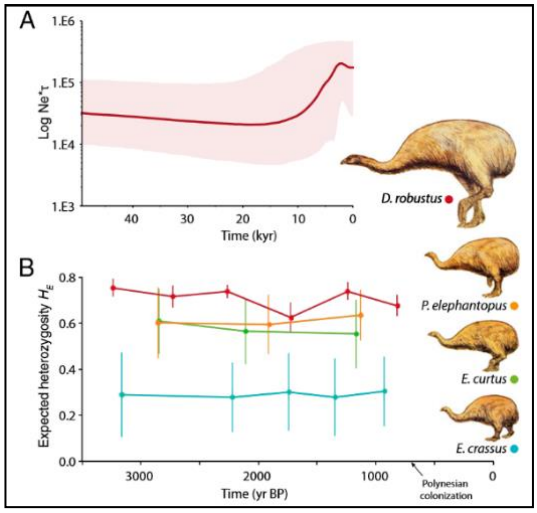

Figure 1.7 Figure from Allentoft et al. (2014) showing the demographic history and genetic diversity of Moa. A) Bayesian skyline plot for *Dinornis robustus (n = 87),* where the y axis depicts the effective female population size multiplied by generation time. Year zero corresponds to the age of the youngest sample at 602 B.P. (B) Expected heterozygosity (HE) for six microsatellite loci, measured across time in the four moa species (n = 188). Data points represent the mean age and mean HE (with SE) of the moa individuals in 1,000-y time bins (Allentoft *et al.*, 2014). If the great auk were unaffected by climatic events and were not in decline prior to the intense hunting that started in ~1500CE, then we could expect to see a result more similar to this shown in the moa, whereby the decline occurred too quickly to be detectable as loss of genetic diversity.

1.4 Aims, objectives & outline of thesis

The principal aim of this thesis is to investigate the evolution and cause of extinction of the great auk, to determine if the great auk was already in decline prior to the period of intense human hunting, and if so, was this due to the impacts associated with climate-driven environmental change, or was it a species that was genetically healthy, and hunting alone caused its demise. This will be achieved through the following objectives:

1. Sample great auk material from museum collections, to establish a dataset that covers as much of their former distribution and over as wide a time scale as possible.
2. Collect and analyse morphometric data from great auk bones to determine the effect of oceanographic zone/geographical location on morphology (as in Burness & Montevecchi, 1992) and use this to develop a hypothesis for our genetic data.

3. Use a palaeogenomic approach to sequence mitochondrial genomes (mitogenome) from the great auk material and use this data to examine temporal changes in levels of genetic diversity throughout the late Pleistocene and Holocene, investigate levels of gene flow between different regions of the great auks' former range, and reconstruct their demographic history to investigate the impact of environmental change on their extinction.
4. Use a palaeogenomic approach to sequence nuclear markers from the great auk samples to provide a more detailed picture of great auk evolution and extinction than using the mitogenome data alone.
5. Explore the mystery of the missing skins from the last two individuals killed in June 1844, by comparing sequences obtained from their organs with those from the candidate specimens proposed by Fuller (1999).
6. Consider the relevance and implications of findings collectively for extant seabird species.

1.5 Chapter overview

Chapter 1: General Introduction

This chapter will provide a detailed overview of literature pertaining to the great auk. It will highlight the gaps in the current knowledge of this species and outline how we aim to address some of the key questions, such as cause of extinction. It also provides an overview of the field of ancient DNA, discussing details of the history, challenges, advancements and uses, in relation to this study.

Chapter 2: Investigating Range Wide Size Variation in Great Auks

In this chapter, we use morphometric data collected from 82 great auk humeri, from sites across their full range, to investigate morphometric variation of great auks in relation to oceanographic region/geographical location. Investigating morphometric variation is relevant to our overall aim of investigating the cause of extinction, as it could indicate population structure, which may reflect lack of gene flow/reproductive isolation, factors which can increase a species risk to extinction. This chapter is currently in preparation for publication.

Chapter 3: Demographic Reconstruction from Ancient DNA Supports Rapid Extinction of the Great Auk

In this chapter we investigate the extinction of the great auk, through the use of 41 mitochondrial genomes. We use this data to determine levels of genetic diversity, population structure and gene flow, across their range and through time. We also present demographic reconstructions to examine the cause of their extinction in relation to climate-driven environmental change vs. hunting. Finally, it uses population viability analysis to assess levels of hunting required to cause extinction. This chapter is currently in preparation for publication.

Chapter 4: Exploring the Possibility of Using Nuclear DNA Sequencing to Characterise the Population Genetics of the Great Auk

As it is generally accepted that a combination of mitochondrial and nuclear DNA should be used to address ecological and evolutionary questions, we attempted to obtain nuclear DNA sequences from a proportion of our great auk samples. Unfortunately, due to low coverage of the targeted markers, we were unable to perform meaningful population genetic analysis. Nevertheless, we make suggestions on how to improve the study for future research.

Chapter 5: An 'Aukward' Tale: A Genetic Approach to Discover the Whereabouts of the Last Great Auks

During the course of this study, we discovered that using the data generated in relation to other chapters, we could attempt to solve one of the mysteries that surrounds the great auk- the whereabouts of the skins from the last two individuals. Using the mitochondrial genomes of the organs from the last documented pair and five candidate specimens, we attempted to find a match between organs and specimen. This chapter is published: *Genes* **2017**, *8*(6), 164; doi:10.3390/genes8060164

Chapter 6- General Discussion

This chapter presents an overall discussion of the key findings from this thesis in relation to published literature. We discuss the findings of our study through comparisons with other extinct birds, perform a hypothetical assessment for the great auks' risk of extinction (if it was extant) and discuss the implications of the findings for related extant seabirds. Finally, we make suggestions for future research questions that will provide further insight into the evolution and extinction of the great auk.

Chapter 2: Investigating Range Wide Size Variation in Great Auks [1]

Tentative author list:

Jessica E. Thomas, Gary R. Carvalho, Michael Hofreiter, M. Thomas P. Gilbert, John R. Stewart [†] and Michael Knapp[†] (Equal contribution [†])

[1] This chapter is in preparation for publication. Thomas, J E., *et al.*, Investigating range wide size variation in great auks

Abstract

It has been documented, for numerous species, that body size variation can exist across its range. This observed variation is often attributed to Bergmann's Rule. Previously published findings suggest that this phenomenon occurs in the extinct great auk (*Pinguinus impennis*). It has been suggested that great auks from Funk Island (North-West Atlantic/Low Arctic oceanographic zone) are larger than those from Scandinavia (North-East Atlantic/Boreal oceanographic zone). This variation has been interpreted as evidence of population structure. Such information is relevant for understanding the extinction of the great auk, as it could suggest limited gene flow between colonies or reproductive isolation, which could potentially have increased the great auks' vulnerability to extinction. Here we investigate size variation in the humerus of great auk, from sites across their former range. Morphometric data was collected from eighty-two great auk humeri, representing individuals from ten countries. The data was split by geographic and oceanographic region. Comparisons were made between North-West Atlantic/Low Arctic oceanographic zone and North-East Atlantic/Boreal oceanographic zone. Additional companions were made between samples from sites regarded as in the vicinity of the Low Arctic oceanographic region/Low Arctic vs. Boreal. The results showed an overall trend of non-statistically significant results. Variation in only one measurement was found to be significant. Our findings contradict the previously published research. The observed overall trend of non-significant results suggests there is no evidence of observable population structure. Therefore, to obtain a better understanding of levels of gene flow and respective population structure/reproductive isolation, we conclude that using genetic data is the most informative course of action.

2.1 Introduction

It is well documented, within vertebrate species, that variation in body size can occur across their range (see studies reviewed in (Meiri & Dayan, 2003)). This phenomenon has been reported in several species of seabirds including the Atlantic puffin (*Fratercula arctica*) (Moen, 1991) and little auk/dovekie (*Alle alle*) (Wojczulanis-Jakubas *et al.*, 2011). The observed variation in size has been attributed to a number of factors, which continue to be discussed and debated. The most commonly cited reason however, pertains to the heat conservation hypothesis, i.e. individuals from colder areas (higher latitude) will be larger than those from warmer climates due to thermoregulation (originally proposed by Bergmann to describe differences at the interspecific level (1847) but was later reformed by Rensch (1938) to refer to within species difference (i.e. Bergmann's rule) (Blackburn *et al.*, 1999; Meiri & Dayan, 2003)). The variation observed may also be explained by the starvation resistance hypothesis (reviewed in (Blackburn *et al.*, 1999) or the availability/quality of food (Moen, 1991; Meiri *et al.*, 2007). While the mechanism of what causes the differences is often debated, the general consensus is that Bergmann's rule is a valid ecological generalisation (Meiri & Dayan, 2003) and one which has also been observed in the extinct great auk (Fig. 2.1) (Burness & Montevecchi, 1992).

Figure 2.1 A mounted great auk from the collections at Royal Belgian Institute of Natural Sciences (RBINS 5355) (Credit Thierry Hubin) and an illustration of a great auk skeleton (originally published Eyton, TC (1869) Supplement to Osteologia Avium. R. Hobsonthis. Image is free from copyright).

For the great auk, such morphometric differences may be relevant for understanding their extinction. It is well documented that the great auk was heavily hunted for its meat, eggs, oil and feathers but it is not known if it was already in decline and vulnerable to extinction, prior to the intense hunting that began ~1500CE (Grieve, 1885; Bengtson, 1984; Fuller, 1999). While morphometric differences along environmental gradients by themselves do not necessarily indicate reproductive isolation (Moen, 1991; Wojczulanis-Jakubas *et al.*, 2015), morphometric variation that is consistent with genetic differences between populations provide strong evidence of reproductive isolation (Hofreiter *et al.*, 2004). If populations were isolated, this may make them more susceptible to extinction under intense hunting pressure.

The great auk was once found across the North Atlantic (Fig. 2.2), inhabiting the waters of multiple oceanographic zones (Burness & Montevecchi, 1992) and across a number of latitudes, therefore groups were likely to experience differences in sea surface and air temperatures. Burness & Montevecchi (1992), compared morphometric data from bones of great auks collected from Funk Island (North-West Atlantic/ Low Arctic oceanographic zone), with individuals from Scandinavia (North-East Atlantic/ Boreal oceanographic zone). They concluded that all bones associated with sub-aqueous flight (coracoid, humerus, ulna and scapula) were significantly larger in great auks from the Low Arctic oceanographic region, than those from the Boreal oceanographic region (Burness & Montevecchi, 1992).

They suggested that the observed differences in size between great auks from the North-West and North-East Atlantic could be due to one of three reasons, or a combination of them. Firstly, genotypically based size differences because of differing environmental pressures (e.g. sea surface temperature), secondly, phenotypically related size differences caused by inter colony differences in prey, and finally, founder effects generated by different sized establishers of different colonies (Burness & Montevecchi, 1992). It is known that many seabird species became fragmented into isolated groups during the last glaciation (Burness & Montevecchi, 1992). Great auk migration is poorly understood (Meldgaard, 1988; Burness & Montevecchi, 1992) and Burness & Montevecchi (1992) suggest that gene flow between colonies of the North-East and North-West Atlantic was unlikely.

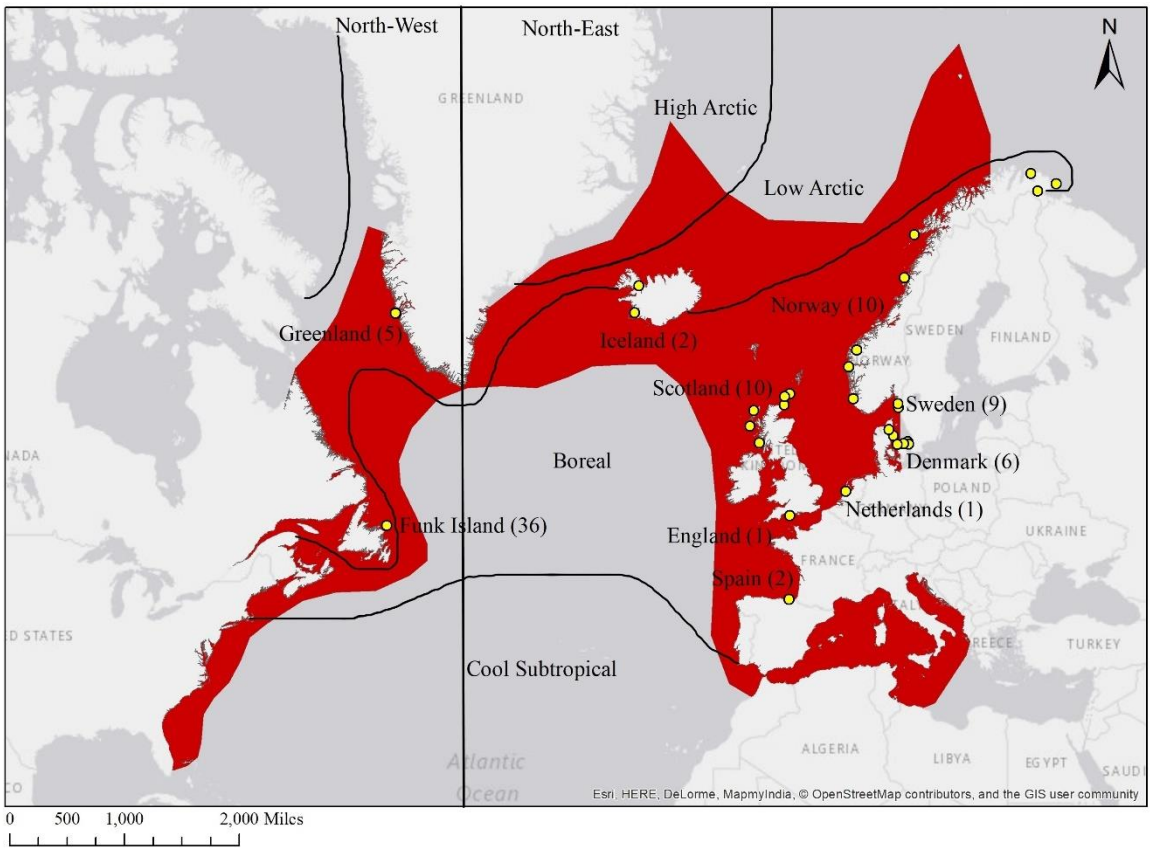

Figure 2.2 The great auks former distribution (red) as defined by the Birdlife International/IUCN (BirdLife International-IUCN, 2016). Sites represented by yellow dots indicate locations from which humeri were collected. Numbers in brackets correspond to number of samples that were used in analysis, from each country. Black lines mark the boundaries of the various oceanographic zones (reproduced from Burness & Montevecchi,1992).

The findings of Burness & Montevecchi (1992) were however, in contrast to those by Hufthammer (1982). Great auks from Scandinavia were found to be smallest in the most northerly locations and from the most recent times (Hufthammer, 1982; Bengtson, 1984; Burness & Montevecchi, 1992). Size differences between groups has also been examined by Meldgaard (1988) who attempted to resolve the origins of wintering great auks off Greenland. When comparing measurements of coracoid, humerus, radius, ulna and femur, no significant difference in size between individuals from Funk Island and Greenland were found. Meldgaard (1988) also found that humerus, radius and ulna measurements of individuals from Iceland fell within the range of those from Funk Island and Greenland (Meldgaard, 1988). Thus, previous morphometric analyses have yielded inconsistent results with regards to whether or not

morphometric evidence for population structure in the great auk exists across the range of the species.

By analysing the most geographically diverse collection of great auk remains to date, this chapter will focus on investigating whether morphometric variation exists in the humerus, between individuals from different regions of the great auks' former geographic range, which could indicate limited gene flow and reproductive isolation of groups.

2.2 Methods

2.2.1 Data collection

Morphometric data was collected from eighty-two great auk humeri (Supplementary Materials Table S2.1, S2.2a&b). Samples represented individuals from the major centres of its former distribution (Fig. 2.2). Measurements were collected using digital callipers following Hufthammer (1982), from points considered to be the most replicable (measurements 'a', 'e', 'f', 'g', 'h', 'i' and 'j' (Fig. 2.3)). All measurements were taken by the same observer (J E Thomas) with the exception of those from the Norwegian sites (excluding Vardø), which were taken by Hufthammer (see Supplementary Materials Table S2.2b).

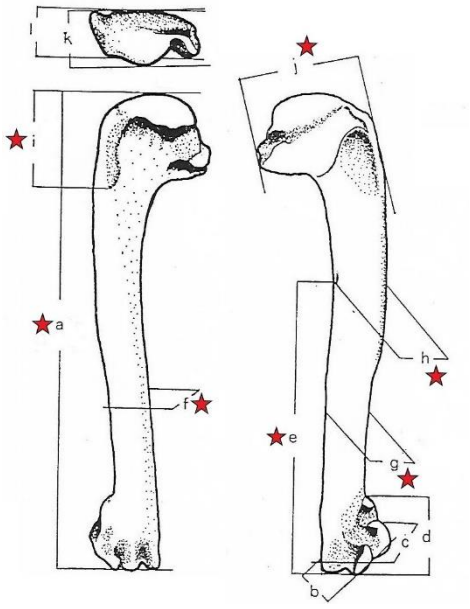

Figure 2.3 Diagram of a great auk humerus, with annotation of measurements by Hufthammer (1982). Measurements used in this study are indicated by a red star. (Details of annotations and measurement guides translated from Hufthammer (1982) can be found in Supplementary Materials Table S2.3, S2.4 and Fig. S2.1).

Whole bones were sampled where possible but due to rarity of samples at many of the sites, this was not always feasible (Fig. 2.4). Consequently, for many of the bones only partial measurement data could be collected. Similarly, several bones showed evidence of damage, such as being worn (Fig. 2.4). Therefore, to ensure maximum accuracy in the dataset, measurements were only collected from points which showed no sign of damage. Where possible, excavation details such as context were used to ensure individual birds were sampled. Where excavation details were not available, only humeri from the right side of the body were measured. For sites where only bones from the left side were available these were measured and the data combined with the that from the right bones as there is no evidence for significant size differences between bones from either side of the body (Burness & Montevecchi, 1992).

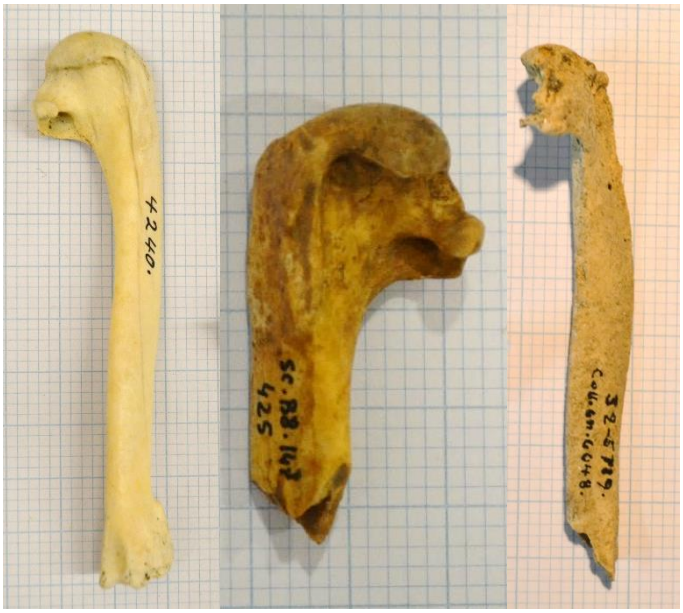

Figure 2.4 Comparisons of whole, broken and damaged humeri to illustrate the various levels of damage to bones encountered when sampling.

2.2.2 Group definitions

Humerus samples were assigned to either the North-East Atlantic (from here in NE) i.e. Boreal oceanographic zone, or the North-West Atlantic (NW) i.e. Low Arctic zone. Samples included in the NE (Boreal) group are those from Iceland, Scotland, England, Denmark, Sweden, Norway, Holland and Spain. The NW (Low Arctic) group includes samples from Funk Island (Canada), and Greenland (Fig. 2.2). A small number of samples from the Boreal group, specifically those from Iversfjord, Nyelv and Vardø, (Norway) and Kollafjarðarnes, (Iceland), are within the 'vicinity' of the Low Arctic zone. To evaluate if the assignment of these samples

influences the results, analyses were conducted with the samples defined as 'vicinity of Low Arctic' assigned to the Boreal group (i.e. the NE vs. NW comparison) and to the Low Arctic group (i.e. Boreal vs. vicinity of Low Arctic/Low Arctic comparison).

2.2.3 Data analysis

Descriptive statistics such as the range, mean, standard deviation (SD) and variance were calculated and bar charts created to visualise mean sizes. Levene's Test for Equality of Variances was used to assess homogeneity of variance between samples of different groups, followed by an Independent Sample *t*-Test to assess mean differences. If the significance value of Levene's Test was greater than 0.05, equal variance was assumed, if it was less than 0.05 then the variability is significantly different and so the results from the 'equal variances not assumed' were reported for the *t*-test. SPSS Statistic v24.0 was used for all statistical analysis for our data (IBM Corp, 2016).

2.2.4 Comparative analysis

As significant size variation in great auks from different oceanographic zones has previously been reported (Burness & Montevecchi, 1992), comparative analysis between our data and that published by Burness & Montevecchi (1992) was performed. Raw measurement data from Burness & Montevecchi (1992) was unavailable, so an Unpaired *t*-Test (performed using GraphPad QuickCalcs Website (GraphPad)) was carried out using the published means and the standard variation, calculated from the published variance.

2.3 Results

2.3.1 General findings

Sample MK86 (115mm), from Funk Island, was found to be the longest bone, in terms of the maximum length measurement (measurement 'a') (Fig. 2.5). The shortest sample, with regards to maximum length, was also a Funk Island sample, MK130 (98.27mm). In terms of width of shaft (which may be of relevance for the phenomenon of dorsoventral flattening, discussed in more detail in 2.4), measurement 'g' (Least dorsoventral corpus width) was the widest for MK79 (11.79mm) and MK93 the narrowest (9.04mm) (both samples from Funk Island). The

dorsoventral corpus width, measurement ‘h’, was widest for sample MK122 (Norway) (14mm) and narrowest for MK104 (Sweden) (9.88mm).

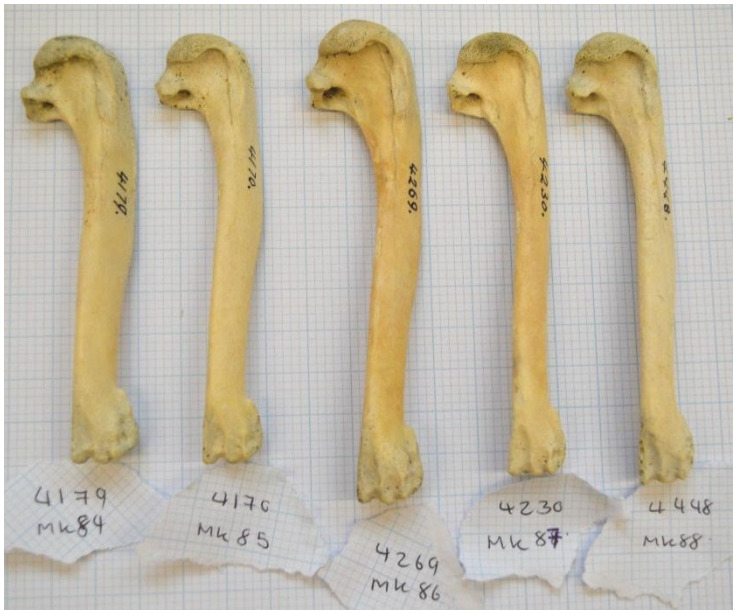

Figure 2.5 Five great auk humeri, collected from Funk Island. MK86 (centre bone) was found to have the greatest maximum length (measurement ‘a’) in our dataset.

2.3.2 North-East vs. North West Atlantic group comparisons

The mean, variance, standard deviation (SD) and range for each of the measurements used in analysis is reported in Table 2.1. Bar graphs were created to visual the mean value of each measurement (Fig. 2.6). Measurements, ‘a’, ‘e’, ‘f’, ‘g’ were larger in the NW group than the NE, and an equal mean measurement was found for measurement ‘i’. For measurements ‘h’ and ‘j’, the mean was larger for the NE group than the NW group (Table 2.1). However, the results of the Independent Samples *t*-Test found only measurement ‘f’ to be statistically significant (<0.05).

		Group		Levene's Test for Equality of Variances		Independent Samples *t*-test for Equality of Means		
		NW	**NE**	***F***	***p***	***t***	**df**	***p* (2-tailed)**
a	Mean	105.64	104.18	0.06	0.81	-1.39	51.00	0.17
	Variance	11.10	11.45					
	SD	3.33	3.38					
	Range	16.73	10.70					
	n	39	14					
e	Mean	61.82	61.11	0.15	0.70	-0.77	47.00	0.44
	Variance	8.50	6.30					
	SD	2.92	2.51					
	Range	11.40	7.33					
	n	36	13					
f	Mean	4.94	4.72	0.58	0.45	-2.55	60.00	0.01
	Variance	0.09	0.12					
	SD	0.29	0.35					
	Range	1.18	1.57					
	n	41	21					
g	Mean	10.48	10.22	1.81	0.18	-1.68	59.00	0.10
	Variance	0.29	0.43					
	SD	0.54	0.66					
	Range	2.75	2.55					
	n	41	20					
h	Mean	11.40	11.51	6.71	0.01	0.60	51.31	0.55
	Variance	0.34	0.87					
	SD	0.58	0.93					
	Range	2.60	4.12					
	n	41	33					
i	Mean	22.57	22.57	0.06	0.81	0.00	68.00	1.00
	Variance	1.05	1.07					
	SD	1.02	1.04					
	Range	4.45	4.06					
	n	39	31					
j	Mean	23.78	24.13	10.98	0.002	0.79	24.16	0.44
	Variance	1.05	3.09					
	SD	1.03	1.76					
	Range	4.56	6.15					
	n	39	19					

Table 2.1 Descriptive statistics for bones from the North-West and North-East Atlantic groups. Independent sample *t*-Test results. Values in red indicate a significant result. SD= Standard deviation, df= Degrees of freedom.

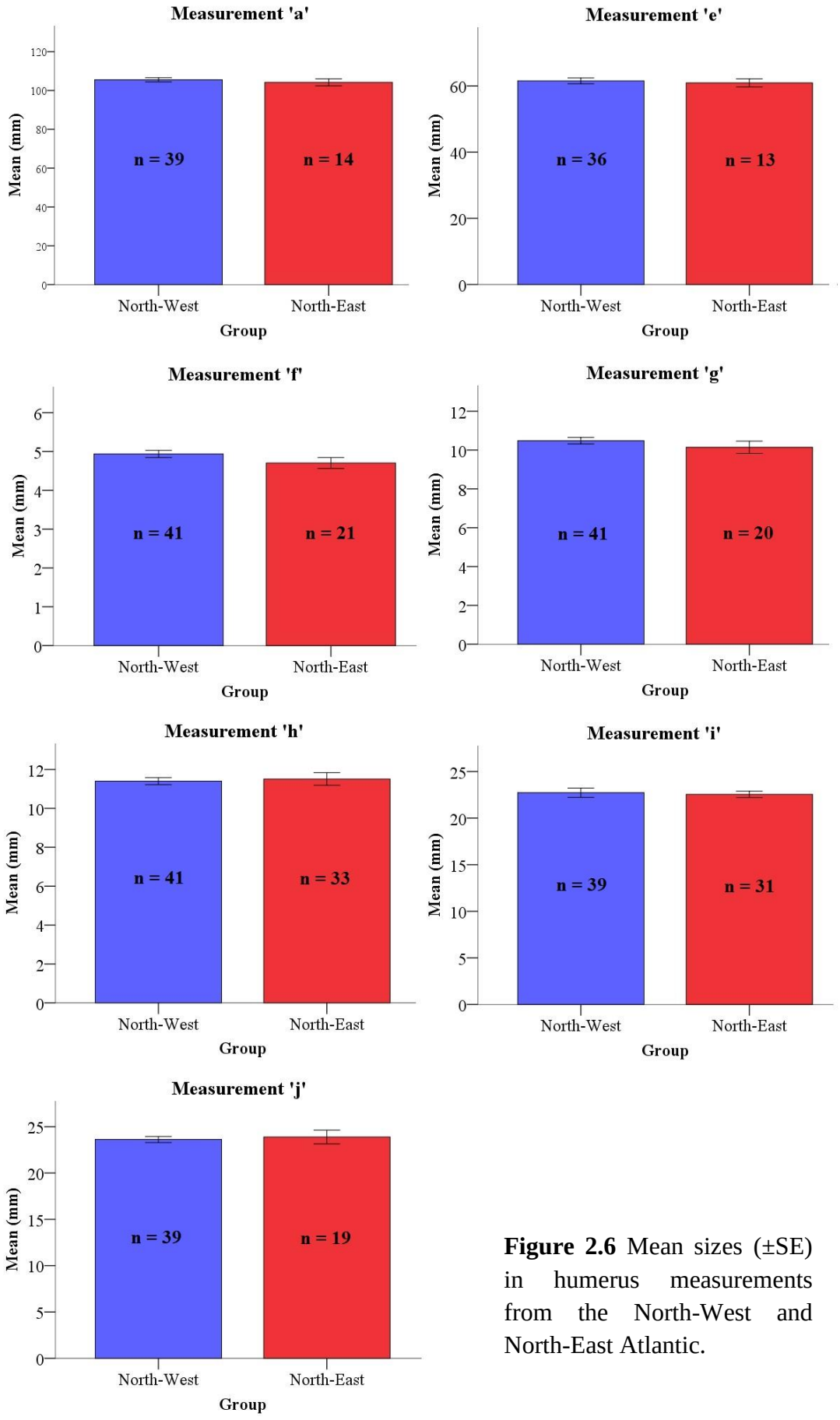

Figure 2.6 Mean sizes (±SE) in humerus measurements from the North-West and North-East Atlantic.

2.3.3 Vicinity of Low Arctic/ Low Arctic vs. Boreal group comparisons

The mean, variance, standard deviation (SD) and range for each of the measurements used in analysis is reported in Table 2.2. Bar graphs were created to visual the mean value of each measurement (Fig. 2.7). Mean results for measurements ‘a’, ‘e’, ‘f’, ‘g’ were larger in the vicinity of Low Arctic/Low Arctic group than that of the Boreal. However, for measurements ‘h’, ‘i’ and ‘j’ the mean was larger for the Boreal group than the vicinity of Low Arctic/Low Arctic group (Table 2.2). The Independent Samples *t*-Test found all differences to be statistically non-significant ($p>0.05$) (Table 2.2).

		Group		Levene's Test for Equality of Variances		Independent Samples *t*-test for Equality of Means		
		Vicinity of Low Arctic/Low Arctic	**Boreal**	***F***	***p***	***t***	**df**	***p* (2-tailed)**
a	Mean	105.71	103.68	0.10	0.75	1.88	51.00	0.07
	Variance	10.67	11.54					
	SD	3.27	3.40					
	Range	16.73	10.70					
	n	41	12					
e	Mean	61.87	60.89	0.14	0.71	1.06	47.00	0.30
	Variance	8.37	6.17					
	SD	2.89	2.48					
	Range	11.40	7.31					
	n	37	12					
f	Mean	4.91	4.75	0.90	0.35	1.75	60.00	0.09
	Variance	0.09	0.14					
	SD	0.30	0.37					
	Range	1.26	1.57					
	n	45	17					
g	Mean	10.46	10.24	1.64	0.20	1.27	59.00	0.21
	Variance	0.31	0.45					
	SD	0.55	0.67					
	Range	2.75	2.55					
	n	44	17					
h	Mean	11.36	11.59	4.15	0.05	-1.17	43.65	0.25
	Variance	0.37	0.86					
	SD	0.61	0.93					
	Range	2.60	4.12					
	n	45	29					
i	Mean	22.52	22.65	0.03	0.87	-0.54	68.00	0.59
	Variance	1.03	1.10					
	SD	1.01	1.05					
	Range	4.45	4.06					
	n	43	27					
j	Mean	23.80	24.13	11.14	0.002	-0.71	20.38	0.49
	Variance	1.06	3.33					
	SD	1.03	1.82					
	Range	4.56	6.15					
	n	41	17					

Table 2.2 Descriptive statistics for bones from the vicinity of Low Arctic/Low Arctic and Boreal groups. Independent sample *t* Test results. Values in red indicate a significant result. SD= Standard deviation, df= Degrees of freedom, value, df=degrees of freedom.

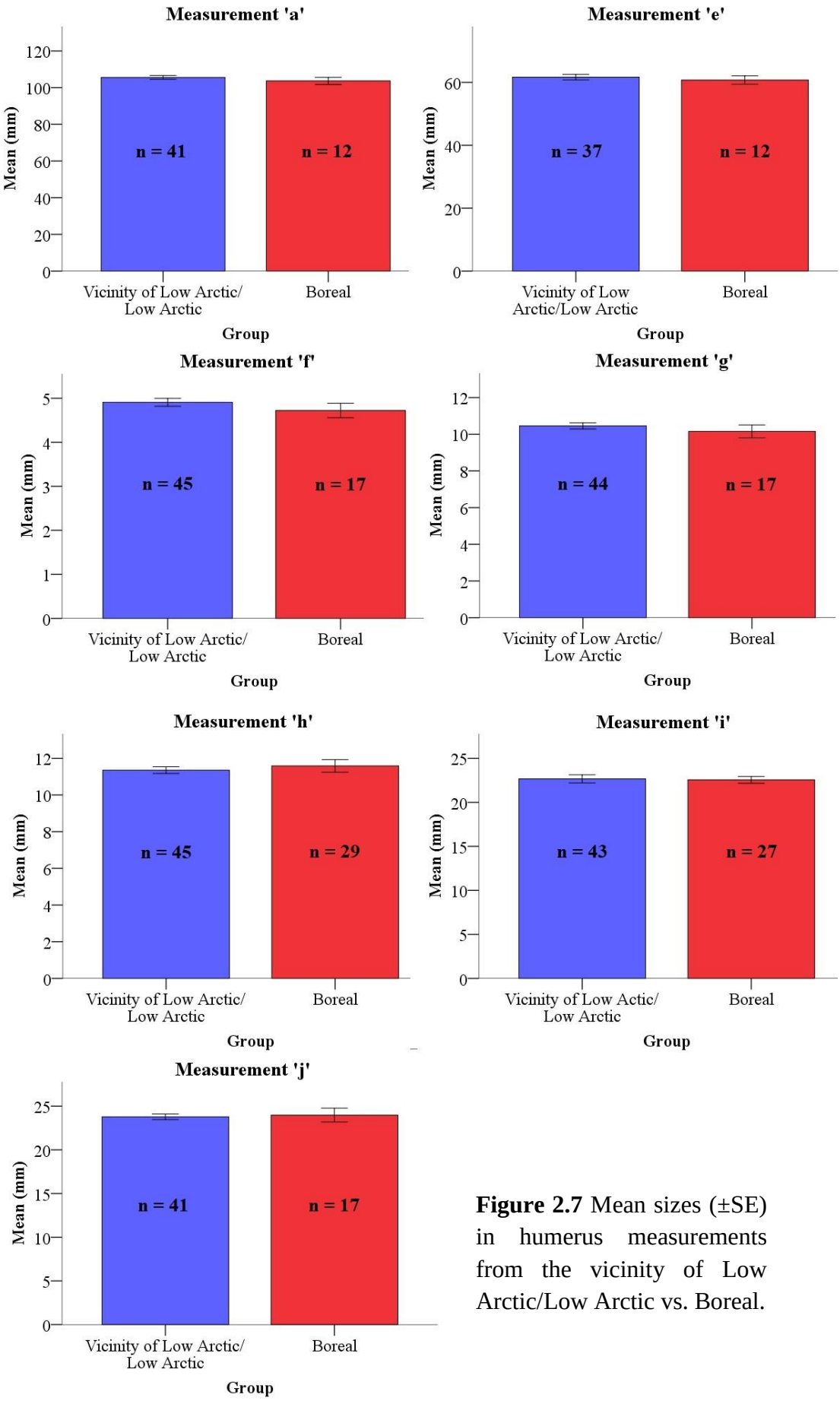

Figure 2.7 Mean sizes (±SE) in humerus measurements from the vicinity of Low Arctic/Low Arctic vs. Boreal.

2.3.4 Comparative analysis

Comparisons between our results from individuals of the NW and the data from Burness & Montevecchi (1992) for the NE, showed that for measurements 'a' and 'g', the NW mean was larger than the NE, however, for measurement 'h', the NE group had a larger mean than the NW. The results of the Unpaired *t*-Test used to compare our data with that of Burness & Montevecchi (1992) showed that all differences found between our NW vs. Burness & Montevecchi (1992) NE were statistically significant (Table 2.3). When we compared our NE results with the NW results published by Burness & Montevecchi (1992), all measurements were larger for the NW group, however, the Unpaired *t*-Test, found all differences to be statistically non-significant (Table 2.3).

	Our NW vs. BM NE			**Our NE vs. BM NW**		
Humerus	***t***	**df**	***p* (2-tailed)**	***t***	**df**	***p* (2-tailed)**
a	6.19	52	< 0.0001	0.38	94.00	0.71
g	3.30	72.00	0.0015	0.47	128.00	0.64
h	4.54	75.00	< 0.0001	0.56	132.00	0.57

Table 2.3 Unpaired *t*-Test results for comparisons between our data, with that published by Burness & Montevecchi (1992). Values in red indicate a statistically significant result. BM= Burness & Montevecchi (1992), df=degrees of freedom.

2.4 Discussion

Population structure in a heavily exploited species, such as the great auk, is a factor that may contribute to the extinction of populations and eventually the whole species. Population structure in great auks was suggested by Burness & Montevecchi (1992), who interpreted morphometric data from both sides of the Atlantic as showing significant size differences between NW (Low Arctic) and NE Atlantic (Boreal) groups, with those from the NW being larger. Our results demonstrate an overall trend of non-significant differences in humeri measurements, between individuals of the NW Atlantic/Low Arctic zone and those of the NE/Boreal zone. Therefore, our results are in contrast to the conclusion drawn by Burness & Montevecchi (1992).

2.4.1 Comparisons in the humerus

The overall result of comparisons between the NW (Low Arctic) vs. NE (Boreal), and vicinity of Low Arctic/Low Arctic vs. Boreal, was that size differences were statistically non-significant. Only one statistically significant result was found in comparisons between NW vs. NE, for measurement 'f', with individuals of the NW being larger. It is important to note that the results of measurements 'h' and 'j', suggest that individuals of the NE/Boreal, were larger than in the NW/Low Arctic. While the differences were found to be non-significant, had sample size for the NE been larger, then this may have affected the result.

The significant size difference found in measurement 'f' and the variation in measurements 'h' and 'j' may be of interest to investigate further in future studies due to the phenomenon known as 'dorsoventral flattening'. Dorsoventral flattening of wing elements has been observed in the flightless Alcidae (*Pinguinus, Mancalla, Praemancalla* and *Alcodes*) (Olson, 1977; Livezey, 1988), although not to the extent as observed in penguins (Elliott *et al.*, 2013). This adaptation is related to wing-propelled diving, i.e. underwater hunting. Measurements 'g', 'h', and 'j', are all dorsoventral shaft measurements and so will reflect this compression of the humerus. As measurement 'f' is a shaft dimension, it will also be affected by the flattening. Therefore, the variations observed in 'f', 'g' 'h' and 'j, while they provide inconsistent results in terms of which group is larger, the variation could be a reflection of environmental adaptions i.e. diving depths required in hunting. This hypothesis is purely speculative and needs to be investigated in much greater detail.

Humeri bones were chosen to be used in this study as they were more prevalent, in comparison to other elements, at the majority of sites. However, it should be discussed that the humerus may not have been the most appropriate choice for investigating size variation in a flightless bird. Wing elements are commonly used in studies investigating size variation (e.g. (Moen, 1991; Wojczulanis-Jakubas *et al.*, 2011, 2015)), but in the great auk, wing elements are reduced in size (in comparison to overall body size) due to being flightless (Livezey, 1988). Hufthammer (1982) suggested that the tarsus length could be a better indicator for size, as larger leg bones would be needed to support a larger/ heavier bird. Alternatively, comparisons of the breadth of the tarsometatarsus may give a better indication for size, as it has been found to correlate with weight in other species (e.g. *Lagopus spp.*) (Stewart, 1999). Interestingly, while Burness & Montevecchi (1992) conclude that they found evidence for significant size

difference between the NW and NE Atlantic, this result was only within bones associated with subaqueous flight. For the leg bones, the differences were all non-significant and the measurement for the tibiotarsus was bigger in the NE group (Burness & Montevecchi, 1992).

2.4.2 Comparative analysis between our study and Burness & Montevecchi (1992)

As Burness & Montevecchi (1992) only reported the result for three of the humerus measurements, we were only able to compare measurements 'a', 'g' and 'h'. Comparative analysis, using Unpaired *t*-Tests, between our data for the NW and that of Burness & Montevecchi (1992) NE, found all differences to be statistically significant. However, measurement 'h' was larger for the NE population of Burness & Montevecchi (1992) than our NW. Therefore, only 2/3 measurements show that individuals of the NW were significantly larger than the NE. When we compared our NE data to Burness & Montevecchi (1992) NW data, despite the measurements of the NW dataset being larger, all differences were statistically non-significant.

2.4.3 Possible explanations for the absence of morphometric variation

As it had previously been hypothesised that oceanographic-related variation existed for the great auk (Burness & Montevecchi, 1992), the overall trend of non-significant variation was unexpected. The difference in result between our study and that of Burness & Montevecchi (1992) could be due to a number of things. Firstly, this study compared samples from sites across the full range of the great auk. Burness & Montevecchi (1992), only compared individuals from Funk Island with those from Scandinavia. While Funk Island is thought to have been the site of the largest breeding colony (Montevecchi and Kirk, 1996), and their presence in Scandinavia is well recorded (Hufthammer, 1982), the range of the great auk is much bigger than just these two areas. As this study sampled individuals over a greater proportion of the former geographic area, this may explain the difference in our results.

While size variation is reported for numerous species of birds (72% of 94 species reviewed in Meiri & Dayan, 2003), it has been found that species which lead a sedentary lifestyle tend to follow Bergmann's Rule more than migratory species (Meiri & Dayan, 2003). Migration in the great auk is poorly understood (Meldgaard, 1988; Burness & Montevecchi, 1992) but it is possible that they made both northward and southward movements aided by oceanic currents

(Bengtson, 1984; Meldgaard, 1988; Burness & Montevecchi, 1992; Serjeantson, 2001). Meiri & Dayan (2003) suggested that birds with a more sedentary lifestyle, which were therefore subjected to natural selection during all seasons, would be more affected by climatic factors than a migratory species which could move. Overwintering sites of the great auk are not as well-known as breeding locations, but there is evidence of wintering great auks off the coast of Greenland (Meldgaard, 1988). While Greenland falls into the Low Arctic oceanographic zone, and sea surface temperatures here are likely to be low, it is possible that the birds did not stay in the same overwintering site all season or even return year on year, choosing to winter in either a Boreal or Low Arctic zone. While of course this is purely speculative, it could be for reason to explain why we do not see size variation in our dataset.

The previously published research by Burness & Montevecchi (1992), suggested that the size differences may have been caused due to absence of gene flow between individuals of either side of the Atlantic. However, the lack of size variation observed in our dataset could be explained by unforeseen migration of the great auk, allowing interbreeding between individuals of either side of the Atlantic. If individuals were able to migrate, therefore allowing gene flow to occur between populations from the North-West and North-East Atlantic then we would be more likely to see the lack of size variation and absence of population structure observed in this dataset than the previously published results, and this may therefore be an alternative explanation to the lack of variation which should be further explored.

2.4.4 What do the results mean for the cause of extinction?

In this study we investigated morphometric differences in the great auk, as it may be relevant for understanding its extinction. As we found an overall trend of non-significant results for morphometric variation, this suggests there is no evidence of observable population structure, between individuals of the NW and NE Atlantic groups. It is important to note that morphometric differences do not always reflect genetic variation (Moen, 1991; Wojczulanis-Jakubas *et al.*, 2015) and vice versa. Therefore, the lack of structure observed in the morphometrics, does not dictate that this reflects the genetic population structure. To fully understand the levels of gene flow and respective population structure/reproductive isolation that could have potentially made the great auk vulnerable to extinction, it stands to order that the most informative course of action is to investigate the species' population genetics.

2.5 Conclusion

To conclude, we found an overall trend of non-significant size variation in the humerus of great auks from either side of the North Atlantic Ocean. However, due to finding one significant result and inconsistencies between which population was larger, inconclusive results prevent a confident conclusion being reached. These results are contradictory to previous published data (Burness & Montevecchi, 1992) and do not provide any evidence for morphometric differences or limited gene flow between populations. The following chapters will therefore focus on molecular data and aim to reconstruct not only population structure, but also demographic changes through time.

2.6 Additional information

This paper is currently in preparation for publication.

2.6.1 Tentative author list

Jessica E. Thomas, Gary R. Carvalho, Michael Hofreiter, M. Thomas P. Gilbert, John R. Stewart [†] and Michael Knapp[†] (Equal contribution [†])

2.6.2 Author contributions

J.E.T., J.R.S., G.R.C., M.H., M.T.P.G., and M.K. conceived the study; J.E.T., and J.R.S., designed the experiments; J.E.T., collected the data (unless otherwise stated); J.E.T., analysed the data; M.H., and J.R.S. provided the initial framework for the study; all authors contributed to writing the paper.

Chapter 2: Supplementary Materials

S2 Tables

Lab ID	Institution	Institution Number (if available)	Site Location & Country	Age (associated period / estimated median age)
MK09	Bournemouth University	Spec No 22722. CH98 (595)	Cladh Hallan, South Uist, Scotland	1135-1035 BC-635-535 BC est. 2850
MK10	Bournemouth University	CH99 (692)	Cladh Hallan, South Uist, Scotland	1135-1035 BC-635-535 BC est. 2850
MK23	University of Southampton	136 (TN86, 012)	Tofts Ness, Sanday, Orkney, Scotland	Neolithic/Early Bronze Age (2000BC) est. 4000
MK24	University of Southampton	293 (TN 0194)	Tofts Ness, Sanday, Orkney, Scotland	Neolithic/Early Bronze Age (2000BC) est. 4000
MK25	University of Southampton	108 (TN86, 0177)	Tofts Ness, Sanday, Orkney, Scotland	Neolithic/Early Bronze Age (2000BC) est. 4000
MK26	Portland Museum	PM1 (Context 189)	Royal Manor Field, Portland, England	Roman age est. 2000
MK30	National Museum of Scotland	NA	Middle Oronsay, Scotland	6200-5100
MK31	National Museum of Scotland	1940-15-40	Knowe of Ramsay, Orkney, Scotland	Neolithic
MK32	National Museum of Scotland	z.2002.175.1 GA753	Freswick Links, Caithness, Scotland	Late Iron Age
MK33	National Museum of Scotland	z.2003.45.1 HP358	Caisteal nan Gillean, Oronsay, Scotland	Mesolithic
MK34	National Museum of Scotland	block 8 F266	Cnip, Lewis, Scotland	Middle Iron Age
MK43	Arkeologi Museoa, Bilboa	SC. B6. L29E.1754	Santa Catalina, Lekeitio, Spain	Magda. Fin. Level III 12900-15000
MK44	Arkeologi Museoa, Bilboa	SC. B8.147.425	Santa Catalina, Lekeitio, Spain	Magda. Fin. Level III 12900-15000
MK48	Natural History Museum of Denmark	10010	Vardø, Norway	Unknown, but unlikely over 5000
MK49	Natural History Museum of Denmark	NA	Vardø, Norway	Unknown, but unlikely over 5000
MK53	Natural History Museum of Denmark	ZMK 4/1861 (19/63)	Mejlgård, Randers, Denmark	Ertebølle culture. 5300 BC – 3950 BC C-14: (*Bos primigenius*) 5115 ± 70 BP
MK54	Natural History Museum of Denmark	ZMK 4/1893 (24/63)	Havnø, Aalborg, Denmark	Ertebølle culture. C-14: 4130 ± 40 BP
MK55	Natural History Museum of Denmark	ZMK 130/1967	Villingebæk Island, Frederiksborg, Denmark	Kongemose Culture, 6000 BC–5200 BC
MK56	Natural History Museum of Denmark	ZMK 101/1951	Reykjavik, Iceland	Unknown

Lab ID	Institution	Institution Number (if available)	Site Location & Country	Age (associated period / estimated median age)
MK57	Natural History Museum of Denmark	ZMK 11/1915	Nivågård, Frederiksborg, Denmark	Early Ertebølleruten Culture
MK58	Natural History Museum of Denmark	ZMK 111/1952	Kangeq, Greenland	Thule Culture 1400-1952 AD
MK59	Natural History Museum of Denmark	ZMK 111/1952	Kangeq, Greenland	Thule Culture 1400-1952 AD
MK62	Natural History Museum of Denmark	ZMK 31/1901	Sølager, Frederiksborg, Denmark	Ertebølleruten Culture, Neolithic Funnel Beaker Culture. C-14: *Phoca groenl*. 5460 ± 40 BP
MK64	Natural History Museum of Denmark	ZMK 104/1968	Olsbjerg, Sejerø, Holbæk, Denmark	Assumed Ertebølleruten Culture
MK67	Natural History Museum of Denmark	ZMK 130/1970	Hope Colony, Godthåbsfjord, Greenland	1721-1728AD 296-289
MK68	Natural History Museum of Denmark	ZMK 130/1970	Hope Colony, Godthåbsfjord, Greenland	1721-1728AD 296-289
MK69	Natural History Museum of Denmark	ZMK 39/1909	Kollafjarðarnes, Iceland	Unknown
MK71	Natural History Museum of Denmark	ZMK 142/1972	Igdlorpait, Greenland	1475 - 1700 century
MK73	University of Amsterdam	NA	Velsen, Netherlands	First Century AD (Roman) 1600-2000
MK74	American Museum of Natural History	Bone # 4547, AMNH 31855	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK75	American Museum of Natural History	Bone # 4408, AMNH 31847	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK76	American Museum of Natural History	Bone # 4218, AMNH 31833	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK77	American Museum of Natural History	Bone # 4387, AMNH 31846	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK78	American Museum of Natural History	Bone # 4624, AMNH 31857	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK79	American Museum of Natural History	Bone # 4339, AMNH 31843	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK80	American Museum of Natural History	Bone # 4549, AMNH 31856	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK81	American Museum of Natural History	Bone # 4446, AMNH 31851	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK82	American Museum of Natural History	Bone # 4204, AMNH 31830	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK81	American Museum of Natural History	Bone # 4446, AMNH 31851	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK82	American Museum of Natural History	Bone # 4204, AMNH 31830	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK83	American Museum of Natural History	Bone # 4285, AMNH 31842	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK84	American Museum of Natural History	Bone # 4179, AMNH 31829	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK85	American Museum of Natural History	Bone # 4170, AMNH 31828	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK86	American Museum of Natural History	Bone # 4269, AMNH 31841	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old

Lab ID	Institution	Institution Number (if available)	Site Location & Country	Age (associated period / estimated median age)
MK87	American Museum of Natural History	Bone # 4230, AMNH 31836	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK88	American Museum of Natural History	Bone # 4448, AMNH 31852	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK89	American Museum of Natural History	Bone # 4534, AMNH 31854	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK90	American Museum of Natural History	Bone # 4526, AMNH 31853	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK91	American Museum of Natural History	Bone # 4423, AMNH 31848	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK92	American Museum of Natural History	Bone # 4425, AMNH 31849	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK93	American Museum of Natural History	Bone # 4429, AMNH 31850	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK94	American Museum of Natural History	Bone # 4211, AMNH 31831	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK95	American Museum of Natural History	Bone # 4214, AMNH 31832	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK96	American Museum of Natural History	Bone # 4219, AMNH 31834	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK97	American Museum of Natural History	Bone # 4223, AMNH 31835	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK98	American Museum of Natural History	Bone # 4232, AMNH 31837	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK99	American Museum of Natural History	Bone # 4240, AMNH 31838	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK100	American Museum of Natural History	Bone # 4254, AMNH 31839	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK101	American Museum of Natural History	Bone # 4264, AMNH 31840	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK102	American Museum of Natural History	Bone # 4374, AMNH 31844	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK103	American Museum of Natural History	Bone # 4380, AMNH 31845	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK104	Göteborgs Naturhistoriska Museum	GNM 6022	Sotenkanalen, Sweden	est. 5000
MK105	Göteborgs Naturhistoriska Museum	GMN 6048	Sotenkanalen, Sweden	est. 5000
MK106	Göteborgs Naturhistoriska Museum	GNM 18684	Nödö, Sweden	est. 5000
MK107	Göteborgs Naturhistoriska Museum	GNM 6112	Nödö, Sweden	est. 5000
MK108	Göteborgs Naturhistoriska Museum	GNM 6849	Skalbank Otterön, Sweden	est. 5000
MK109	Göteborgs Naturhistoriska Museum	GNM 6194	Sotenkanalen, Sweden	est. 5000
MK110	Göteborgs Naturhistoriska Museum	GNM 18,685	Sotenkanalen, Sweden	est. 5000
MK111	Göteborgs Naturhistoriska Museum	GNM 9644	Skalbank Otterön, Sweden	est. 5000
MK114	Göteborgs Naturhistoriska Museum	GNM 18,692	Sotenkanalen, Sweden	est. 5000

Lab ID	Institution	Institution Number (if available)	Site Location & Country	Age (associated period / estimated median age)
MK115	University of Bergen	JS. 603	Iversfjord, Finnmark, Norway	3000-2400BC, est. 5016-4416
MK116	University of Bergen	JS.523	Storbåthelleren, Nordland, Norway	3300-2790/790-100BC, est. 5316-2116
MK117	University of Bergen	JS.528	Storbåthelleren, Nordland, Norway	3300-2790/790-100BC, est. 5316-2116
MK118	University of Bergen	JS.412	Viste, Rogaland, Norway	Mesolithic 7000-5000BC
MK120	University of Bergen	JS.427	Skjonghelleren, Møre og Romsdal, Norway	Older Iron Age 500BC-500AD
MK121	University of Bergen	JS.371	Nyelv, Finnmark, Norway	5000-4000BP C-14: 4160+- 80
MK122	University of Bergen	JS.260	Kirkehelleren, Nordland, Norway	Mesolithic to 500AD
MK124	University of Bergen	JS.445	Gronehelleren, Sogn og Fjordane, Norway	Late Mesolithic- recent times
MK125	Zoological Museum Oslo	1213 (1)	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK126	Zoological Museum Oslo	1213 (2)	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK127	Zoological Museum Oslo	1213 (3)	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK128	Zoological Museum Oslo	1213 (4)	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK129	Zoological Museum Oslo	1213 (5)	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old
MK130	Zoological Museum Oslo	1213 (6)	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000 years old

Table S2.1 Sample information, including institute where samples were collected and institution number if assigned, site information from where sample was found and any associated age or date information that was available.

		Measurement (mm) (/ indicates no data)						
Lab ID	**Country**	**a**	**e**	**f**	**g**	**h**	**i**	**j**
MK58	Greenland	/	/	4.51	10.82	11.26	22.32	23.35
MK59	Greenland	104.2	63.72	4.43	9.47	10.79	25.46	22.74
MK67	Greenland	/	58.65	4.42	9.57	10.32	21.67	/
MK68	Greenland	109.06	63.59	4.38	10.17	11.49	23.77	24.1
MK71	Greenland	105.47	/	4.9	10.23	11.8	21.99	24.29
MK74	Canada	103.98	59.13	5.15	10.36	11.29	21.12	25.93
MK75	Canada	100.71	57.77	4.83	9.96	10.9	22.54	23.36
MK76	Canada	103.69	62.99	5.19	11.23	11.99	21.87	25.56
MK77	Canada	107.64	60.04	5.15	10.96	11.08	23.25	23.52
MK78	Canada	110.36	61.76	5.05	11.16	12.15	24.21	24.52
MK79	Canada	110.46	61.99	5.56	11.79	11.87	23.16	24.08
MK80	Canada	107.33	62.58	4.43	9.74	11.1	21.16	23.87
MK81	Canada	105.62	56.84	4.83	10.33	11.08	22.54	23.87
MK82	Canada	103.58	65	4.94	10.5	11.4	21.43	23.02
MK83	Canada	101.94	66.84	4.39	10.16	10.21	21.78	22.26
MK84	Canada	101.09	57.98	5.03	10.54	11.85	21.8	23.98
MK85	Canada	102	60.78	5.02	9.91	10.96	22.5	22.53
MK86	Canada	115	67.77	4.9	10.81	12.5	24.01	26.48
MK87	Canada	107.04	63.16	4.76	9.9	10.79	22.03	23.98
MK88	Canada	109.99	67.1	5.28	10.76	11.57	22.86	24.21
MK89	Canada	104.31	64.93	5.04	10.58	11.76	21.05	24.45
MK90	Canada	106.55	63.15	4.99	10.99	11.84	21.58	23.83
MK91	Canada	110.05	62.71	5.07	10.17	11.74	21.01	24.76

		Measurement (mm) (/ indicates no data)						
Lab ID	**Country**	**a**	**e**	**f**	**g**	**h**	**i**	**j**
MK92	Canada	105.1	/	5	10.72	11.15	/	23.5
MK93	Canada	106.86	63.36	4.77	9.04	10.56	23.09	22.47
MK94	Canada	108.43	65.21	5.27	11.33	11.65	23.63	22.3
MK95	Canada	104	/	4.9	10.62	11.42	22.59	23.67
MK96	Canada	105.55	59.39	5.02	10.32	11.82	23.07	23.77
MK97	Canada	104.33	60.3	5.14	10.06	11.15	23.98	22.15
MK98	Canada	103.06	60.26	4.76	10.39	10.96	22.47	23.75
MK99	Canada	102.26	61.66	4.67	10.08	10.1	22.78	22.71
MK100	Canada	109.72	61.46	5.2	11.08	12.38	24.06	24.23
MK101	Canada	105.5	62.61	4.77	10.76	11.7	22.19	/
MK102	Canada	102.66	/	5.12	10.57	11.46	21.01	23.52
MK103	Canada	105.97	58.39	4.96	10.85	11.65	/	22.89
MK125	Canada	108.28	62.15	4.72	10.5	11.52	23.05	23.93
MK126	Canada	107.38	65.25	4.87	10.38	11.47	22.08	25.05
MK127	Canada	107.08	60.74	5.18	11.18	12.7	22.7	25.09
MK128	Canada	102.73	56.37	4.89	10.72	10.8	23.57	21.92
MK129	Canada	102.53	62.27	5.5	10.6	11.27	22.58	24.5
MK130	Canada	98.27	57.5	5.4	10.5	11.8	22.17	23.47

Table S2.2a Raw data for samples in the North-West Atlantic/ Low Arctic oceanographic zone. Measurement data collected by J. Thomas. Measurements correspond to labels on Fig. 2.3.

		Measurement (mm) (/ indicates no data)						
Lab ID	**Country**	**a**	**e**	**f**	**g**	**h**	**i**	**j**
MK09	Scotland	/	57.94	/	9.5	/	/	/
MK10	Scotland	/	/	/	9.65	10.96	/	/
MK23	Scotland	/	/	/	/	11.77	23.28	/
MK24	Scotland	100.74	58.42	5.05	9.8	12.23	22.13	/
MK25	Scotland	104.00	63.77	4.98	10.57	12.07	23.67	25.69
MK26	England	/	/	/	/	12.55	24.08	24.02
MK30	Scotland	/	/	/	/	11.2	23.56	/
MK31	Scotland	/	/	/	/	/	21.46	22.26
MK32	Scotland	109.09	61.51	5.59	11.02	11.93	23.72	25.81
MK33	Scotland	/	/	/	/	/	23.8	24.87
MK34	Scotland	106.06	62.86	4.62	10.63	12.21	22.08	24.48
MK43	Spain	/	/	/	/	/	22.75	/
MK44	Spain	/	/	/	/	/	23.7	/
MK48	Norway	106.8	/	4.61	10.47	10.58	23.21	/
MK49	Norway	107.65	63.79	4.85	10.5	10.72	21.92	25.18
MK53	Denmark	/	/	/	/	11.57	22.5	/
MK54	Denmark	/	/	/	/	10	/	/
MK55	Denmark	107.96	62.95	4.93	10.98	12.23	22.9	24.48
MK56	Iceland	102.07	58.17	4.89	9.96	11.48	22.35	/
MK57	Denmark	/	/	/	/	11.3	/	/
MK62	Denmark	/	/	/	/	/	23.84	25.46
MK64	Denmark	/	/	/	/	/	22.62	/
MK69	Iceland	/	/	4.3	9.24	10.32	21.49	/
MK73	Netherlands	106.2	62.7	5.17	11.75	13.39	22.49	25.89

		Measurement (mm) (/ indicates no data)						
Lab ID	**Country**	**a**	**e**	**f**	**g**	**h**	**i**	**j**
MK104	Sweden	/	/	/	/	9.88	21.55	20.15
MK105	Sweden	/	/	4.43	9.92	10.48	20.7	/
MK106	Sweden	/	/	/	10.09	11.09	22.68	/
MK107	Sweden	104.16	62.11	4.26	10.31	11.08	24.56	25.08
MK108	Sweden	99.17	56.46	4.59	10.72	11.51	22.35	21.92
MK109	Sweden	98.39	60.76	4.87	9.94	11.33	21.07	/
MK110	Sweden	/	/	4.36	9.47	10.85	21.63	21.9
MK111	Sweden	104.98	63	4.02	10.62	10.92	22.5	22.34
MK114	Sweden	101.28	/	4.66	9.2	10.06	/	/
MK115 (AKH)	Norway	/	/	4.7	/	12.2	/	/
MK116 (AKH)	Norway	/	/	4.6	/	11.3	/	/
MK117 (AKH)	Norway	/	/	/	/	12.2	/	/
MK118 (AKH)	Norway	/	/	4.9	/	12.2	/	/
MK120 (AKH)	Norway	/	/	/	/	12.2	/	/
MK121 (AKH)	Norway	/	/	/	/	/	21.4	23
MK122 (AKH)	Norway	/	/	/	/	14	23.1	26.3
MK124 (AKH)	Norway	/	/	4.8	/	12	20.5	26.1

Table S2.2b Raw data for samples in the North-East Atlantic/Boreal oceanographic zone. Measurement data collected by J. Thomas, with the exception of those labelled (AKH) which were collected by A-K. Hufthammer, University of Bergen. Measurements correspond to labels on Fig. 2.3. Samples in yellow are those in the 'vicinity' of Low Arctic zone.

Measurement ID	Name	Description
a	Maximum length	Measure from the Caput humeri (proximal)- Epicondylus ulnaris (distal)
e	Length from Foramen corpus humeri - Epicondylus ulnaris	Measure from the caudal apex Epicondylus ulnaris
f	Least lateromedial corpus (shaft) width	Move the slider to find the smallest measurement
g	Least dorsoventral corpus width	Move the slider to find the smallest measurement
h	Dorsoventral corpus width	Measured 'on' the Foramen corpus humeri
i	Length of Caput humeri - impression of Latissimus dorsi muscle	Measure the cranial apex Caput humeri - caudal end of the impression of the Latissimus dorsi muscle
j	Dorsoventral proximal width	Rest callipers on Crista tuberculi dorsalis (dorsal)- Apex tuberculi ventralis (ventral)

Table S2.3 Measurement ID and description on how measurements were collected. Translated from Hufthammer (1982).

NB. Measurement points used in this study correspond to the following in Burness & Montevecchi (BM) (1992) Humerus: ‘a’ = ‘a’ BM, ‘h’ = ‘b’ BM, ‘g’ = ‘c’ B

Number	Description
1	Caput humeri
2	Crista tuberculi dorsalis
3	Sulcus transversus
4	Incisura collaris
5	Foramen pneumatic
6	Fossa tricipitalis
7	Tuberculum dorsale
8	Tuberculum intermedium
9	Condylus radialis
10	Condylus ulnaris
11	Epicondylus ulnaris
12	Fossa olecrani
13	Ectepicondylis radialis
14	Foramen corpus humeri
15	Entepicondylus radialis
16	Tuberculum pronator brevis
17	Impressio musculi latissimus dorsi
18	Apex tuberculi ventralis
19	Facies musculi bicipitis

Table S2.4 Numbers corresponding to labels in Fig. S2.1 (Hufthammer, 1982).

S2 Figures

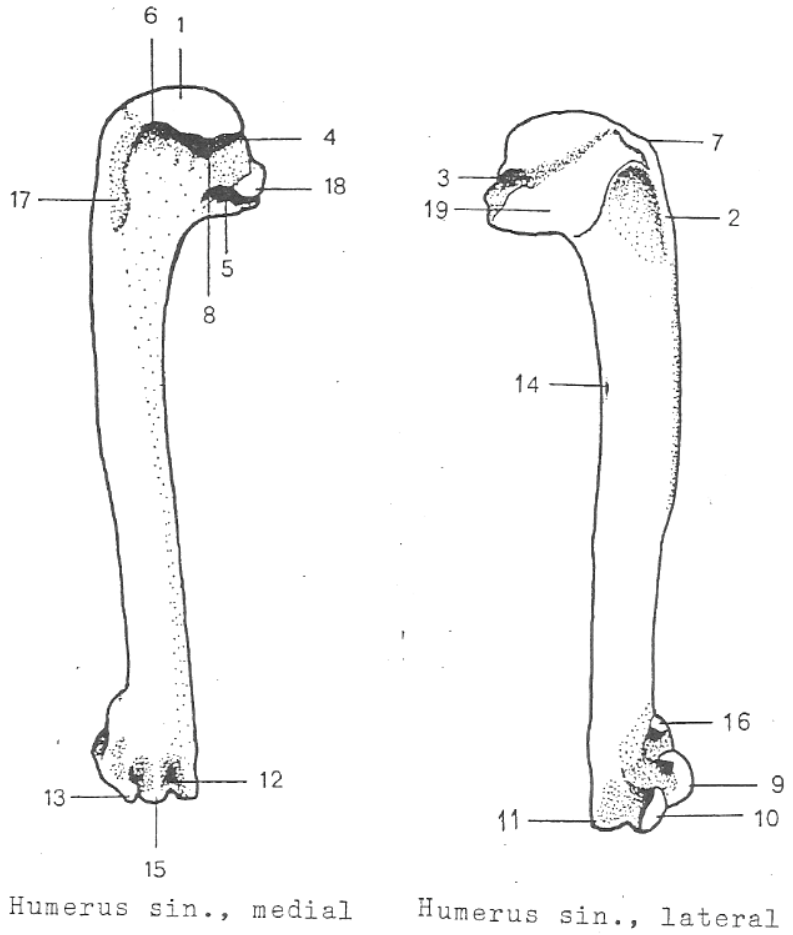

Figure S2.1 Diagram of great auk humerus (Hufthammer, 1982) with numbers corresponding to landmarks used in taking measurements described in Table S2.3. Numbers correspond to those in Table S2.4.

Chapter 3: Demographic Reconstruction from Ancient DNA Supports Rapid Extinction of the Great Auk[2]

Tentative author list:

Jessica E. Thomas, Gary R. Carvalho[†], James Haile, Nicolas J. Rawlence, Michael D. Martin, Simon Y. W. Ho, Arnór Þ. Sigfússon, Vigfús A. Jósefsson, Morten Frederiksen, Jannie F. Linnebjerg, Jose A. Samaniego Castruita, Jonas Niemann, Mikkel-Holger S. Sinding, Marcela Sandoval-Velasco, André E. R. Soares, Christina Barilaro, Julia Best, Dirk Brandis, Chiara Cavallo, Mikelo Elorza, Kimball L. Garrett, Maaike Groot, Friederike Johansson, Jan T. Lifjeld, Göran Nilson, Dale Serjeanston, Paul Sweet, Errol Fuller, Anne Karin Hufthammer, Morten Meldgaard, Jon Fjeldså, Beth Shapiro, Michael Hofreiter, John R. Stewart[†], M. Thomas P. Gilbert[†] and Michael Knapp[†] ([†]Equal contribution)

[2] This chapter is in preparation for publication. Thomas, J E., *et al.*, Demographic reconstruction from ancient DNA supports rapid extinction of the great auk

Abstract

The great auk, (*Pinguinus impennis*), was a flightless bird, once abundant and widely distributed across the North Atlantic. However, it was heavily exploited for its eggs, meat, oil and feathers, and during its final years of existence, as a display item in natural history exhibitions. Despite wide scientific and public interest in the species, it remains unclear if hunting alone was responsible for its demise, or whether it was already in decline due to climate-driven environmental change. As one of few marine bird species in the Northern Hemisphere to have gone extinct in the late Holocene, the great auk provides an excellent example for investigating extinction risk due to environmental change and hunting. To investigate the relative importance of climate-driven environmental change and human hunting, we employed ancient DNA analyses and high-throughput sequencing to generate complete mitochondrial genomes from 41 great auks from across their Holocene and Late Pleistocene North Atlantic range. Intriguingly, our data shows significant levels of genetic diversity and gene flow persisting through time, across the range of the species. A lack of genetic structure is supported by data collected by collaborators from GPS-equipped drifting capsules released in the North Atlantic Ocean, which demonstrate how great auks would have been able to move between colonies, aided by ocean currents, wind and waves. Demographic reconstructions reveal it also had a large and stable effective population size, with no evidence of decline or recent genetic bottlenecks, suggesting that its extinction happened too rapidly to shape its detectable genetic diversity. Finally, we used population viability analysis to estimate what level of hunting would have been required to drive its extinction in a period of fewer than 350 years. This revealed that harvesting 5-7% of the population caused extinction in a significant number of simulation replications, depending on population size. Our findings are consistent with the current consensus that human hunting was the primary cause of the great auks' extinction.

3.1 Introduction

In the early 19th century, the French palaeontologist Georges Cuvier recognised species extinction as a natural process and essential part of life on earth (Cuvier, 1813). Human activity has arguably accelerated the rate of species extinction (Barnosky *et al.*, 2011; Ceballos *et al.*, 2015; Ceballos *et al.*, 2017), and obtaining a better understanding of the causes of and contributions to species loss and extinction risk is now an essential premise for conservation biology (Owens & Bennett, 2000; Purvis *et al.*, 2000). The study of past extinction events can provide valuable insights into the drivers and dynamics of global biodiversity (Halliday, 1978), and recent extinctions in particular have the potential to highlight factors contributing to species extinction that are highly relevant to the conservation of extant endangered species. As one of few marine bird species in the Northern Hemisphere to have gone extinct in the Holocene, the great auk (*Pinguinus impennis*) is an exemplar of a recent extinction. Studying its demise can help to understand the extinction risk from environmental change and hunting, and yield insights for the conservation management of extant species.

The great auk (Fig. 3.1), was a large, flightless black and white bird thought to have once numbered in the millions (Birkhead, 1993). It was a member of the order Charadriiformes, within the Alcidae family, with its closest extant relative being the razorbill (*Alca torda*) (Moum *et al.*, 2002). Distributed around the North Atlantic (Fig. 3.1), colonies could be found along the east coast of North America, especially the island of Newfoundland. The great auk also bred on islands off Iceland and Scotland, and was found throughout Scandinavia (Norway, Denmark and Sweden), southward through western Europe, down to the Mediterranean (Grieve, 1885; Fuller, 1999). Despite populations being found around the North Atlantic, there is little information about migration patterns or movement between colonies. It has also been suggested that its population also underwent northward and southward movements aided by oceanic currents (Bengtson, 1984; Meldgaard, 1988). Furthermore, in addition to suggestions of separate populations in Europe in prehistoric times (Bengtson, 1984), studies of great auks from different oceanographic zones revealed oceanographic-related size variation, with individuals from the North-West/Low Arctic oceanographic zone, being larger than their conspecifics from the Boreal waters of Scandinavia (Burness & Montevecchi, 1992). It has also been proposed that there was limited or no interbreeding between the colonies on either side of the North Atlantic (Burness & Montevecchi, 1992), which could cause high levels of population genetic structure.

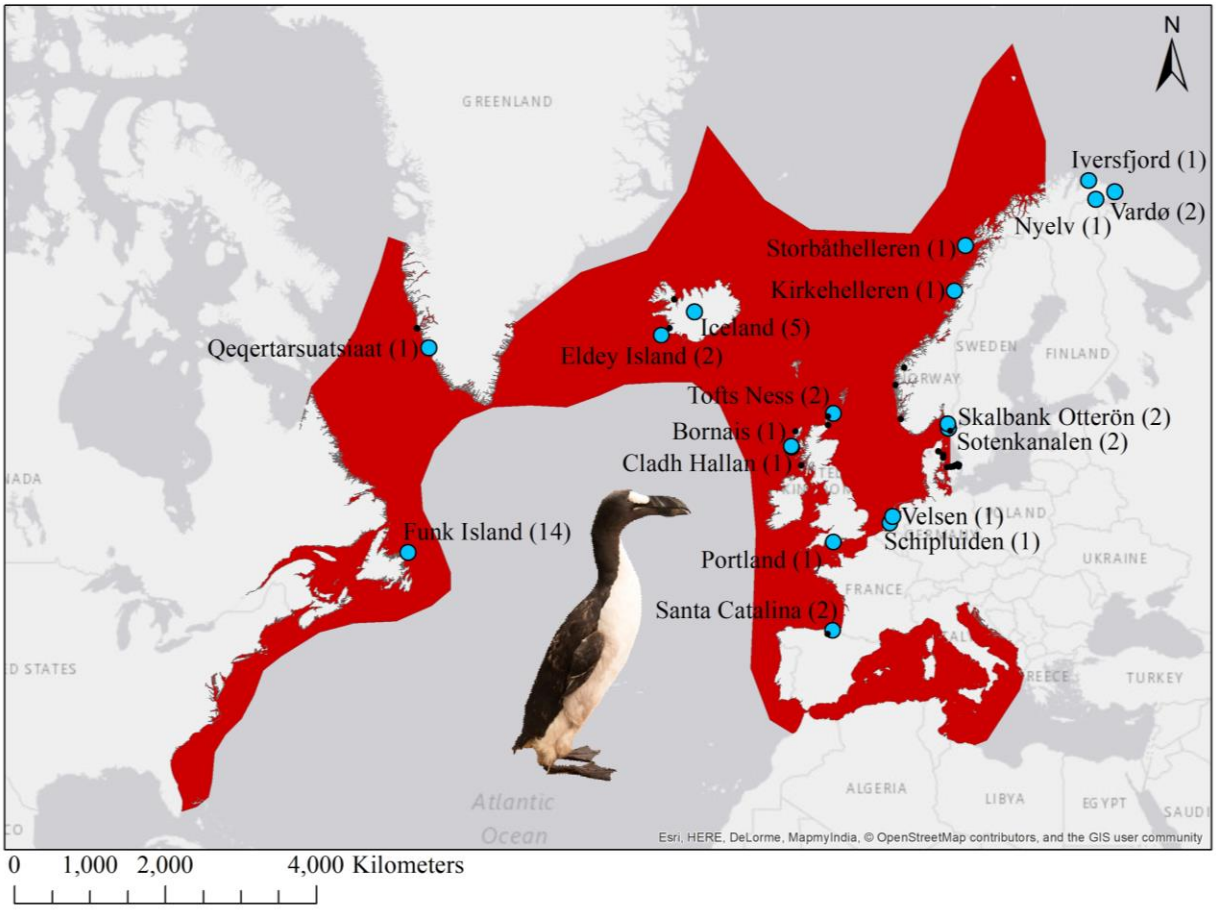

Figure 3.1 The great auk (centre) and its former distribution as defined by BirdLife International/IUCN (BirdLife International-IUCN, 2016). Sites marked with blue dots represent samples used in present analysis, with location and number of sample from that site in brackets. Sites from which material was sampled but that was not sequenced or did not pass filtering settings to be included in further analysis are in black and unlabelled.

Great auks spent the vast majority of their time at sea, only coming ashore to breed for a few weeks during the summer. On land, the great auk was clumsy and vulnerable to both predators and humans, rendering it an easy target for exploitation. It is well documented by both the archaeological and historical records, that the great auk was hunted by humans throughout its existence. In prehistoric times, they were hunted for their meat and eggs by Beothuk Indians in North America (Fuller, 1999; Gaskell, 2000), the Inuit of Greenland (Meldgaard, 1988), Scandinavians, Icelanders (Bengtson, 1984), Magdalenian hunter-gathers in the Bay of Biscay (Laroulandie *et al.*, 2016) and possibly even Neanderthals on the eastern side of their range (Halliday, 1978). However, it was around 1500CE that intensive hunting began by European seamen visiting the fishing grounds of Newfoundland (Steenstrup, 1855; Bengtson, 1984; Fuller, 1999; Gaskell, 2000). Initially, like the local people, the sailors killed the great auk for

sustenance, albeit at a greater intensity. However, the development of commercial hunting for the feather trade towards the end of the 1700s, led to a much higher exploitation level, which, according to contemporary examples provided in the literature, was performed in a more wasteful and destructive manner (Kirkham & Montevecchi, 1982; Fuller, 1999; Gaskell, 2000). Several authors have suggested that it was hunting for the feather trade that caused their extinction rather than hunting for their meat and eggs (Kirkham & Montevecchi, 1982; Gaskell, 2000; Pope, 2009). Towards the end of their existence, as their rarity increased, preserved great auks became desirable for private and institutional collections. The last breeding pair ever reliably seen were killed on Eldey Island, Iceland, June 1844 for this reason (Steenstrup, 1855; Newton, 1861; Grieve, 1885; Bengtson, 1984; Fuller, 1999; Gaskell, 2000; Thomas *et al.*, 2017), although, it is likely that they would have been killed for food, regardless of the collection trade. Records of individuals observed after 1844 have been discussed in the literature (e.g. 1848 in Vardø, Norway (Newton, 1861; Fuller, 1999) and 1852 in Newfoundland (Newton, 1861; Grieve, 1885; Fuller, 1999)) and BirdLife International/ IUCN recognises the last sighting as 1852 (BirdLife International, 2016b), however, there is some degree of uncertainty about these later sightings (Grieve, 1885; Fuller, 1999).

It is, however, not known if the species was already in decline and at risk of extinction prior to the period of intensive hunting, or if climate-driven environmental change had also influenced their demise (Bengtson, 1984; Birkhead, 1993; Fuller, 1999). Here, we use palaeogenetic data to reconstruct the great auks' population structure and population dynamics through time, in order to gain insights into population dynamics as well as the timing of their demise. Specifically, we use these data to explore the likely factors contributing to extinction, including potential effects of hunting and environmental change. Furthermore, we combine our genetic results with data collected from GPS-equipped drifting capsules deployed in the North Atlantic, in order to obtain a better understanding of possible great auk migration routes, thus contextualize our population genetic structure results. Lastly, we model the rate of great auk harvesting that would have been required to drive them to extinction and consider implications for extant seabirds exposed to hunting and impacts of climate change.

3.2 Methods

3.2.1 Sample collection

Great auk material for use in ancient DNA (aDNA) extraction, was sourced from various institutions. Samples were chosen to represent individuals from the major centres of its former geographic distribution, over as great a time period as possible (Supplementary Materials Table S3.1). Bones were sampled via drilling using a Dremel 107 2.4mm engraving cutter to obtain powdered bone (Fig. 3.2), or using a Dremel cutting wheel, which allowed removal of sections of bones which were later powdered using a sonic dismembrator. Tissue or feather samples from mounted specimens and tissue samples from the organs of the last pair collected in 1844 were sampled and processed using methods appropriate to sample type, as described in Thomas et al. (2017).

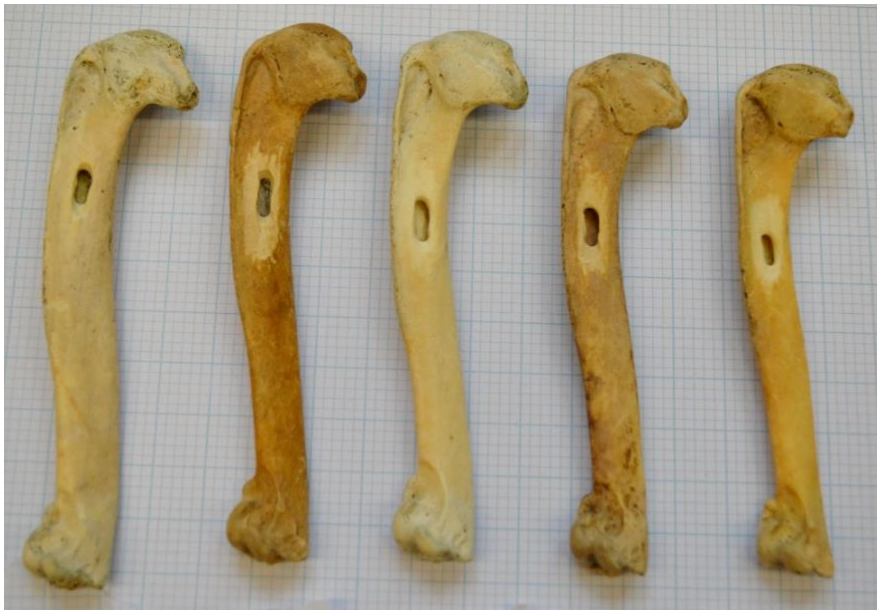

Figure 3.2 Great auk humeri, collected from Funk Island, following sampling to collect bone powder for use in DNA extraction (bones part of the collection at the American Museum of Natural History) (Credit: J. Thomas).

3.2.2 DNA extraction

All laboratory work prior to Polymerase Chain Reaction (PCR) amplification was carried out in the designated aDNA laboratories of the Natural History Museum of Denmark and the University of Otago. Strict aDNA protocols were followed to avoid contamination of the experiments. For each DNA extraction and library build, no-template controls were used to

test for contamination by exogenous DNA. All post-PCR work on amplified DNA was carried out in separate laboratory facilities (Knapp *et al.*, 2012).

Genomic DNA was extracted from 20-60mg bone powder using the method described by Dabney et al. (2013). In short, the bone powder was digested using an EDTA based extraction buffer and DNA purified using a Qiagen MinElute column. After washing with ethanol-based wash buffers (Qiagen), the DNA was eluted in TE buffer for storage. Full details of laboratory protocols can be found in Appendix 1.

3.2.3 Data generation

To determine the most appropriate and effective methods for the samples in this study, a preliminary experiment was performed in which we compared library build methods (Double-stranded vs. Single-stranded), as well as the effectiveness of using hybridisation enrichment capture for whole mitochondrial genomes versus shotgun sequencing. Details of the method comparisons can be found in Supplementary Materials Additional Information S3.1. The results of the comparisons were used to determine the best combination of methods which was then used for the remaining samples in the study and was performed as follows.

Single-stranded sequencing libraries were prepared from aDNA extracts following the protocol by Gansauge & Meyer (2013), with modifications as described in Bennet et al. (2014). For some samples, double stranded libraries were also built using the protocol described by Meyer & Kircher (2010) (Supplementary Materials Table S3.2). Hybridisation capture was used to enrich libraries for great auk mitochondrial DNA following the MYcroarray MYbaits Sequence Enrichment protocol v2.3.1 (MYcroarray MYbaits, 2014). 100mer mitochondrial DNA baits (MYcroarray MYbaits) with 50bp tiling were designed using a hybrid reference mitochondrial genome (mitogenome), that was constructed using the killdeer (*Charadrius vociferus*) mitogenome (assembled from whole genome data; BioProject: PRJNA212867 (Zhang, B. Li, *et al.*, 2014)), with orthologous gene regions replaced by those of great auk where available (GenBank: AJ242685), and those from the razorbill (*Alca torda*) (GenBank: AJ301680, EF380281, EF380318, X73916) when great auk data were unavailable (Supplementary Materials Fig. S3.1). Samples were sequenced on Illumina platforms (HiSeq2500 and MiSeq (further details in Supplementary Materials Table

S3.2)) at the Danish National High-Throughput DNA Sequencing Centre or by New Zealand Genomics Limited.

3.2.4 Read processing and consensus sequence filtering

Demultiplexing of raw sequence data was performed by the respective sequencing centres. Read processing of demultiplexed sequence data was performed as described in Thomas et al. (2017) using the PALEOMIX v1.2.5 pipeline (Schubert *et al.*, 2014) and included software tools to remove adapters, filter bases based on quality (AdapterRemoval v2.1.7 (Lindgreen, 2012; Schubert *et al.*, 2016)), and map reads to the reference mitogenome (Burrows-Wheeler Aligner v0.5.10, (Li and Durbin, 2009)). At the time of these analyses, a great auk mitogenome had been published (GenBank: KU158188.1 (Anmarkrud & Lifjeld, 2017)), and was thus, available for the mapping assembly of our mitogenomes rather than mapping against the composite mitogenome used for bait design (see above). PCR duplicates were removed using MarkDuplicates within Picard v1.8.2 (Broad Institute) and the rmdup function within SAMtools (Li *et al.*, 2009). The Genome Analysis Toolkit (GATK) v3.6.0 was used to correct for misaligned reads to the reference mitogenome using the RealignerTargetCreator and IndelRealigner functions (McKenna *et al.*, 2010). Finally, MapDamage2 (Jonsson *et al.*, 2013) was employed to rescale base quality scores according to their probability of being damaged, thereby removing residual aDNA damage patterns, and the UnifiedGenotyper algorithm within GATK v3.6.0 was used to determine haploid genotypes for individual samples.

Consensus sequences were produced using two filtering criteria, following Chang et al. (2017). This allowed us to evaluate and control for potential errors associated with missing or low coverage data caused by sequencing error and damage-derived substitutions which are characteristic of aDNA samples. For 'relaxed' consensus calling, the per-individual read depth was set to only include bases with a minimum of 3-fold coverage. Bases called for the consensus sequence had to be present at a frequency higher than 33%. To be included in the final alignment, no more than 33% of bases could be missing from the consensus sequence of an individual compared to the reference sequence. All bases not meeting these criteria were called as 'N'. For the 'strict' settings, the per-individual read depth for called bases was set to include only bases with at least 10-fold coverage. Geneious v-10.1.3 (Kearse *et al.*, 2012) was used to filter bases so that the majority base had to be present in more than 90% of reads. For an individual to be included in the final alignment, a threshold of no more than 20% missing

sites from an individual's consensus sequence was applied. All analyses were conducted with both the 'relaxed' and 'strict' alignments. As the results obtained from analyses of both alignments yielded consistent results, only the analyses of the relaxed alignment, which included more individuals, will be discussed here. Additional details for analysis using the 'strict' alignment is provided in the Supplementary Materials Additional Information S3.2.

3.2.5 Sequence alignment and model selection

Following read processing, data were aligned using Seaview v4.0 (Gouy *et al.*, 2010) with the algorithm *Muscle -maxiters2 -diags.* The alignment was manually checked for errors using BioEdit v7.2.5 (Hall, 1999), and Tablet v1.16.09.06 (Milne *et al.*, 2013) was used to view the rescaled Binary Alignment Map (BAM) file for each sample. Model selection was performed using JModelTest v2.1.10 (Darriba *et al.*, 2012) to determine the most suitable nucleotide substitution model. The best model suggested by the Bayesian Information Criterion (BIC) was Hasegawa-Kishino-Yano (HKY) (Hasegawa *et al.*, 1985) + invariable sites (+I). The best model suggested by the Akaike Information Criterion (AIC) was HKY (Hasegawa *et al.*, 1985) +I + gamma distribution (+G). We therefore selected the BIC suggested model, as this had the least parameters.

3.2.6 Network analyses

Great auk population structure was reconstructed using a Median Joining Network (Bandelt *et al.*, 1999) as implemented in PopART (Leigh & Bryant, 2015). Genetic diversity through space and time was visualised using statistical parsimony and a temporal haplotype network as implemented in TempNet (Prost & Anderson, 2011). Initially, age categories were chosen based on changes in climate and hunting pressure. Samples were divided into four groups: over 12,000 years old (i.e. Late Pleistocene samples), >1,000 – 12,000 years old (i.e. Holocene samples when hunting pressure was low and opportunistic), ~500 years old (i.e. the period when intense hunting began but diversity should be representative of the previous 12,000 years) and finally, those less than 250 years old (i.e. samples from during the period of intense hunting, including samples from the last pair ever reliably seen killed in 1844). Alternative age categories tested can be found in Supplementary Materials Additional Information S3.3. For samples with date information available this was used to determine age group. For the 16 samples without date information, BEAST v1.7.5 and v1.8.4 (Drummond *et al.*, 2012) were

used to estimate their age based on evolutionary rate and the known age of other samples (i.e. information used as tip calibrations), allowing us to place them in the appropriate category.

3.2.7 Demographic inference using unpartitioned data set

We reconstructed changes in population size over time in two ways. Firstly, we performed analysis on the unpartitioned data set using the BEAST v1.7.5 and v1.8.4 (Drummond *et al.*, 2012) software packages. We used a uniform prior of 0.002 - 0.003 substitutions per site per million years (subst/site/Myrs) as an estimate of the mitochondrial substitution rate of great auks. This estimate was based on body mass-corrected substitution rates for bird mtDNA (Nabholz *et al.*, 2016). Specifically, our range of possible rates incorporates those estimated for the ancient murrelet (*Synthliboramphus antiquus*) (0.0024 subst/site/Myrs), which is the closest relative to the great auk in the published data set, and the Adélie penguin (*Pygoscelis adeliae*) (0.002 subst/site/Myrs)), which has a similar weight and life history to that of the great auk (Nabholz *et al.*, 2016). To refine the rate estimate, we also incorporated the age of individual samples (tip calibration) as prior information. For samples with a stratigraphic age estimate, we assigned the median age of the stratigraphic estimate. For samples without any age information, we sampled the age from a uniform prior, in which the upper and lower limits were constrained to reasonable time limits based on known sample history. For example, Funk Island samples were excavated from the upper layers, therefore they are most likely to be from individuals killed in the hunting since 1500CE, thus we applied an upper limit of 1000 ybp and lower limit of 0, as we know they are extinct. Calibration information for each sample with a known age is provided in Table S3.3 Supplementary Materials. All analyses were conducted with and without tip calibrations to evaluate the influence of the tip age estimates on the analyses.

The best fitting evolutionary rate and coalescent priors were identified using Bayes Factor model comparisons (Kass & Raftery, 1995) as implemented in Tracer v1.5.0 (Rambaut *et al.*, 2009). We compared strict vs. uncorrelated relaxed lognormal clock and, independently, a constant size coalescent tree prior vs. an exponential growth tree prior and vs. the more flexible Bayesian Skyline Plot (BSP) (with 10 groups). In all cases, in order to reduce error introduced by more parameter rich models, these more complex models (relaxed clock; BSP) were only accepted if they were ‘decisively’ better than the respective simpler models (Kass & Raftery, 1995).

For each BEAST analysis, we ran three independent Markov Chain Monte Carlo (MCMC) chains for 20 million generations, sampling trees and model parameters every 2,000 generations, with the first 10% of each chain discarded as burn-in. We compared the results from each run in Tracer v1.5.0 to confirm convergence of the MCMC chains and achieving at least 200 as effective sample size (ESS) for each parameter (Drummond and Bouckaert, 2015). We then combined the posterior samples from the three independent chains and estimated model parameter values. We calculated the maximum clade credibility (MCC) tree using TreeAnnotator v1.8.4 (part of the BEAST software package Drummond *et al.*, 2012). Trees were visualised in FigTree v1.4.3 (Rambaut, 2016).

3.2.8 Analysis using partitioned dataset

3.2.8.1 Molecular dating

To infer the evolutionary rate and timescale, we performed a Bayesian phylogenetic analysis of the mitogenome sequences from the 25 dated samples. The 'relaxed' sequence alignment was analysed using BEAST 1.8.4 (Drummond *et al.*, 2012). The evolutionary timescale was estimated using a strict clock model, with the sampling times of the mitogenomes serving as calibrations for the clock (Rambaut, 2000). A uniform prior of (10^{-10},10^{-4}) was used for the mutation rate, with a separate rate assigned to each subset of the data defined by the partitioning scheme.

Six partitioning schemes were compared for the data, varying in the degree of partitioning and the resulting number of data subsets (Table 3.1). For each data subset, the best-fitting model of nucleotide substitution was selected using the Bayesian information criterion in Modelgenerator (Keane *et al.*, 2006). Constant-size and exponential-growth coalescent tree priors were also compared for the data. Analyses using a skyride coalescent prior (Minin *et al.*, 2008) were attempted but invariably failed to converge, which strongly suggested overparameterization.

Partitioning scheme[a]	Marginal likelihood[b]	
	Constant size	**Exponential growth**
Unpartitioned	-24151.6	-24143.6
2 subsets: (CR rRNA tRNA) (PC1 PC2 PC3)	-24222.3	-24212.4
3 subsets: (CR) (rRNA tRNA) (PC1 PC2 PC3)	-24162.4	-24150.1
4 subsets: (CR) (rRNA tRNA) (PC1 PC2) (PC3)	-23659.7	-23647.5
5 subsets: (CR) (rRNA tRNA) (PC1) (PC2) (PC3)	-23248.7	-23235.9
6 subsets: (CR) (rRNA) (tRNA) (PC1) (PC2) (PC3)	-23229.1	-23217.5

[a]Components of the mitogenome are the ribosomal RNA genes (rRNA), transfer RNA genes (tRNA), three codon positions of the protein-coding genes (PC1, PC2, and PC3), and the control region (CR). [b]Marginal likelihoods were estimated by stepping-stone sampling with 25 path steps, each with a chain length of 2,000,000 steps.

Table 3.1 Marginal likelihoods of four partitioning schemes and two tree priors for the 25 dated mitogenomes.

The marginal likelihood was computed for each combination of partitioning scheme and tree prior, using stepping-stone sampling with 25 path samples (Xie *et al.*, 2011). A partitioning scheme with six data subsets provided the best fit to the data. This approach is consistent with previous analyses of time-structured mitogenomic data sets (e.g., (Anijalg *et al.*, in press). The exponential-growth coalescent tree prior yielded higher marginal likelihoods than the constant-size coalescent tree prior, but the posterior distribution of the population growth rate was highly right-skewed with a mode very close to zero. Therefore, we used the constant-size coalescent tree prior for our analysis.

Posterior distributions of parameters were estimated by Markov chain Monte Carlo (MCMC) sampling. Samples were drawn every 5000 steps from a chain with a total length of 50,000,000 steps. The MCMC analysis was run in duplicate to check for convergence and the first 10% of samples were discarded as burn-in. Effective sample sizes of the model parameters were estimated to ensure that they were all over 200, which indicates sufficient sampling.

To test for the presence of temporal structure in the data set, we performed a date-randomization test (Ramsden *et al.*, 2008). We estimated mutation rates from 20 replicate data sets in which the sampling times were permuted and compared these with the rate estimate from the original data set. Two different criteria can be used to determine whether the data set has sufficient temporal structure for generating a reliable estimate of the mutation rate (Duchêne *et al.*, 2015): if the mean or median estimate from the original data set is not contained within the 95% credibility intervals of the rate estimates from the date-randomised

replicates (CR1), or if the 95% credibility intervals of the rate estimates from the date-randomised replicates do not overlap with the 95% credibility interval of the rate estimate from the original data set (CR2).

A second analysis was performed in BEAST, in which the 16 undated mitogenomes were included in the data set. A uniform prior of either (0,1000) or (0,5000) was specified for the ages of these mitogenomes, depending on independent information about the context of the samples (Shapiro *et al.*, 2011). All other settings and priors matched those used in the analysis of the 25 dated samples.

For comparison, we used two additional methods to estimate the mutation rate. First, we used TempEst (Rambaut *et al.*, 2016) to estimate the mutation rate using regression of root-to-tip distances against sampling times. Second, we analysed the data using least-squares dating in LSD (To *et al.*, 2016). For both of these methods, a phylogram was estimated from the dated mitogenome sequences using maximum likelihood in RAxML 8 (Stamatakis, 2014). Rooting of the tree was inferred by maximising the R-squared value in TempEst and by minimising the objective function in LSD.

3.2.8.2 Phylogenetic analysis

To infer the relationships among all of the mitogenomes, including the undated sequences, we performed a maximum-likelihood analysis using RAxML. The data set was partitioned according to the best-fitting scheme identified in our Bayesian analyses, as described above. Bootstrap support values were estimated using 1000 replicates.

3.2.9 Extinction simulation using population viability analysis software

We employed the population viability analysis (PVA) software Vortex10.2.8.0 (Lacy and Pollak, 2014) in order to obtain an estimate of the number of great auk individuals hunted annually, and the rate at which a given intensity of hunting would result in population collapse and extinction. Scenarios were set to run for a 350-year period, as intense hunting began in ~1500CE (Steenstrup, 1855; Bengtson, 1984; Fuller, 1999; Gaskell, 2000) and no confirmed sightings of great auks occurred later than 1852 (Grieve, 1885; Fuller, 1999; BirdLife International, 2016b). Data produced in this study shows a lack of significant population

structuring of the great auk (see Fig. 3.3 & 3.4), and we therefore consider great auks of the North Atlantic to comprise a single panmictic population. Scenarios were run as a population-based model. Models were also run under various extinction definition scenarios to check the impact this had on results (extinction defined as: only 1 sex remains, Total N < critical size 50, Total N < critical size 500) (The '50:500 rule' refers to a species risk of extinction as defined by Franklin (1980)) . With little information known about great auk life history, many inputs for the model derive from estimates based on information from closest living relatives. Life history parameters were estimated based on the judgements by Bengtson (1984) with considerations of the great auks' larger size and flightlessness taken into consideration. Mortality rates for the various life stages were estimated using information on survival rates of extant species described in De Santo & Nelson, (1995). Full details of the parameters used in the model with related justification can be found in Supplementary Materials Additional Information S3.4.

With no reliable method of inferring census size from the effective population size, we tested three initial population sizes that seemed appropriate based on our mean effective population size inferred from BEAST analysis and estimates from the literature regarding great auk population size (Contemporary accounts describe finding 'infinite' and 'innumerable' flocks on Funk Island (Gaskell, 2000), which based on numbers of common murres (*Uria aalge*) that breed there today, and taking into account the great auks larger size, are estimated to be in excess of 100,000 breeding pairs (Birkhead, 1993)). The population sizes tested were: one million, three million and ten million. To estimate hunting pressure, models were run in which various proportions of the population were harvested to investigate the effect that harvest percentage had on the extinction rate, rather than the actual number taken. Initially 2%, 5% and 10% were tested and found to be too low (no extinction, or low levels of extinction) or too high (extinction occurred too quickly), therefore we tested intermediate levels of 6% and 7%. For a population size of 1 million, we also tested harvest rates of 3% and 4%. Within each of the total population harvest percentage tested, harvest for each age group was calculated as follows. The age categories for harvest rate are 0-1, 1-2, 2-3, 3-4, 4-5 and over 5 for both males and females. For over 5, we estimate this to be 75% of the harvest number, with the remaining 25% spilt between the other four age categories (1-5). There are less juvenile bones present in the fossil record and only one museum mount is assumed to be an immature bird; however, young are known to have been used as bait. As we know eggs were collected, we allowed for this in the model. Egg removal levels tested were 50,000 or 100,000, which based on estimates

of breeding pairs at Funk Island (>100,000) (Birkhead, 1993), allowed for us to test the impact of at least half of the eggs laid annually being harvested, and if 'all' eggs laid annually on Funk Island were harvested.

3.2.10 Tracking migration routes using GPS capsules

In order to obtain a better understanding of how feasible it would be for great auks to move between colonies of the North Atlantic, two GPS-equipped drifting capsules were released from a helicopter around 40km southeast of the Reykjanes peninsula on 10.01.2016. The capsules contained a North Star TrackPack GPS tracking device (https://www.northstarst.com/asset-trackers/trackpack/), which uploaded precise location data 6 times a day for up to 2 years, through the GlobalStar satellite network. The device was built into a capsule designed by Verkís Consulting Engineers in Iceland. The housing consisted of two cylinders made of transparent plexi-glass. The inner cylinder contained the GPS device, which had to point upwards to send the signal. The outer cylinder contained glycerol in which the inner cylinder floated. The ends of the outer cylinder were inside a soft buoy to prevent it from breaking if washed ashore.

3.3 Results

3.3.1 DNA extraction and sequencing

In total, DNA libraries were sequenced from 74 samples. Following sequence read processing, 41 samples passed the relaxed filtering settings and were included in downstream analysis. The mean average number of aligned bases to the reference mitogenome for all samples used in the analysis was 55.12bp, with a range of 41.21-86.95bp. Unique coverage of these samples ranged from 6.39x to 430.09x with average coverage of 72.5x (see Supplementary Materials Table S3.4 for full PALEOMIX summary stats and Supplementary Materials Additional Information S3.5 for details of Molecular Preservation). Following removal of all sites that contained gaps in at least one sequence, the relaxed alignment was 9994 base pairs (bp) in length. The strict alignment contained 25 individuals with a sequence length of 9373 bp. GenBank Accession numbers for samples used in analysis can be found in Supplementary Materials Table S3.5.

3.3.2 Genetic diversity and phylogeographic analyses

Haplotype diversity was high, with only two of 41 individuals showing identical haplotypes in the final relaxed alignment. When split into different age groups, each age group was characterised by a completely different set of haplotypes and no reduction of haplotype diversity could be identified in younger samples (Fig. 3.3). The statistical parsimony network (Fig. 3.3) and median joining network (Fig. 3.4) show no phylogeographic or temporal structure in the distribution of haplotypes.

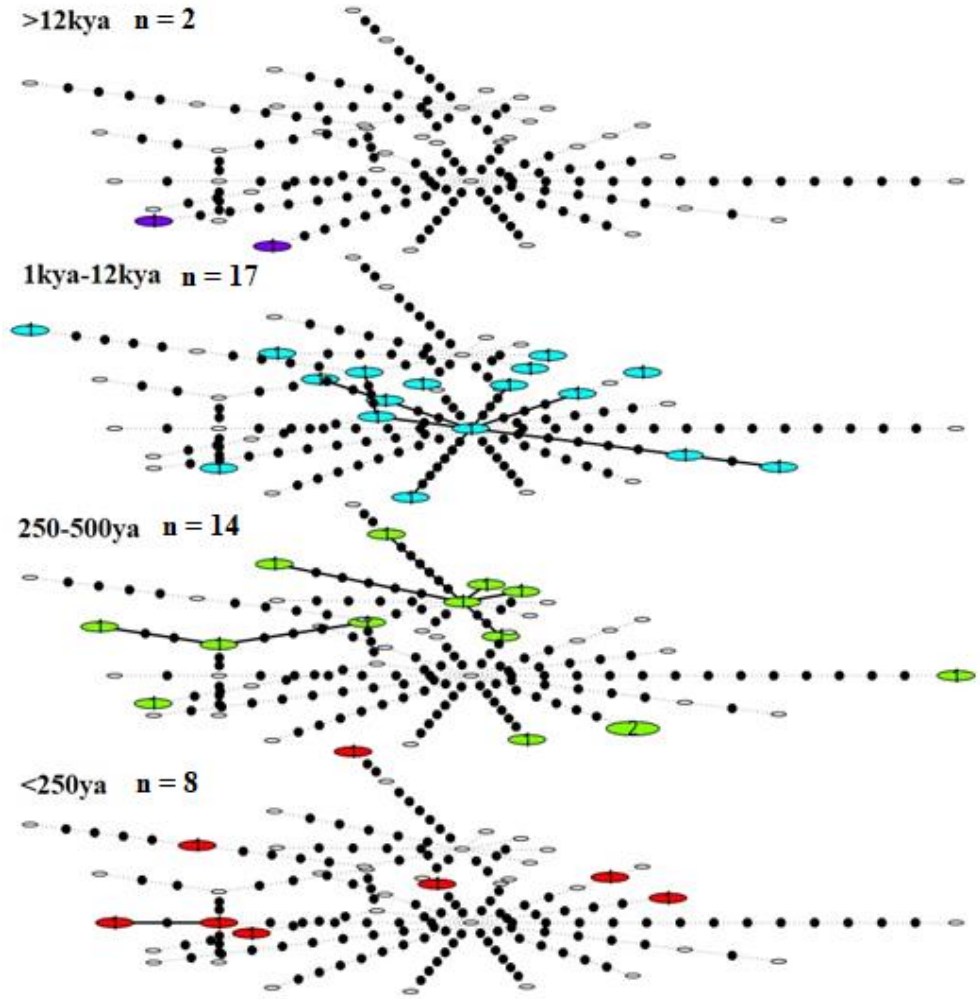

Figure 3.3 TempNet showing the haplotype diversity through time. Samples were split into four age categories. Each age category is represented by the same statistical parsimony network, with haplotypes present in this time category shown in colour, haplotypes present in another time category shown as empty dots, and mutation between haplotypes marked as filled dots. All samples have been included in this figure. For samples which did not have dates, BEAST was used to estimate their date and they were entered into the appropriate category.

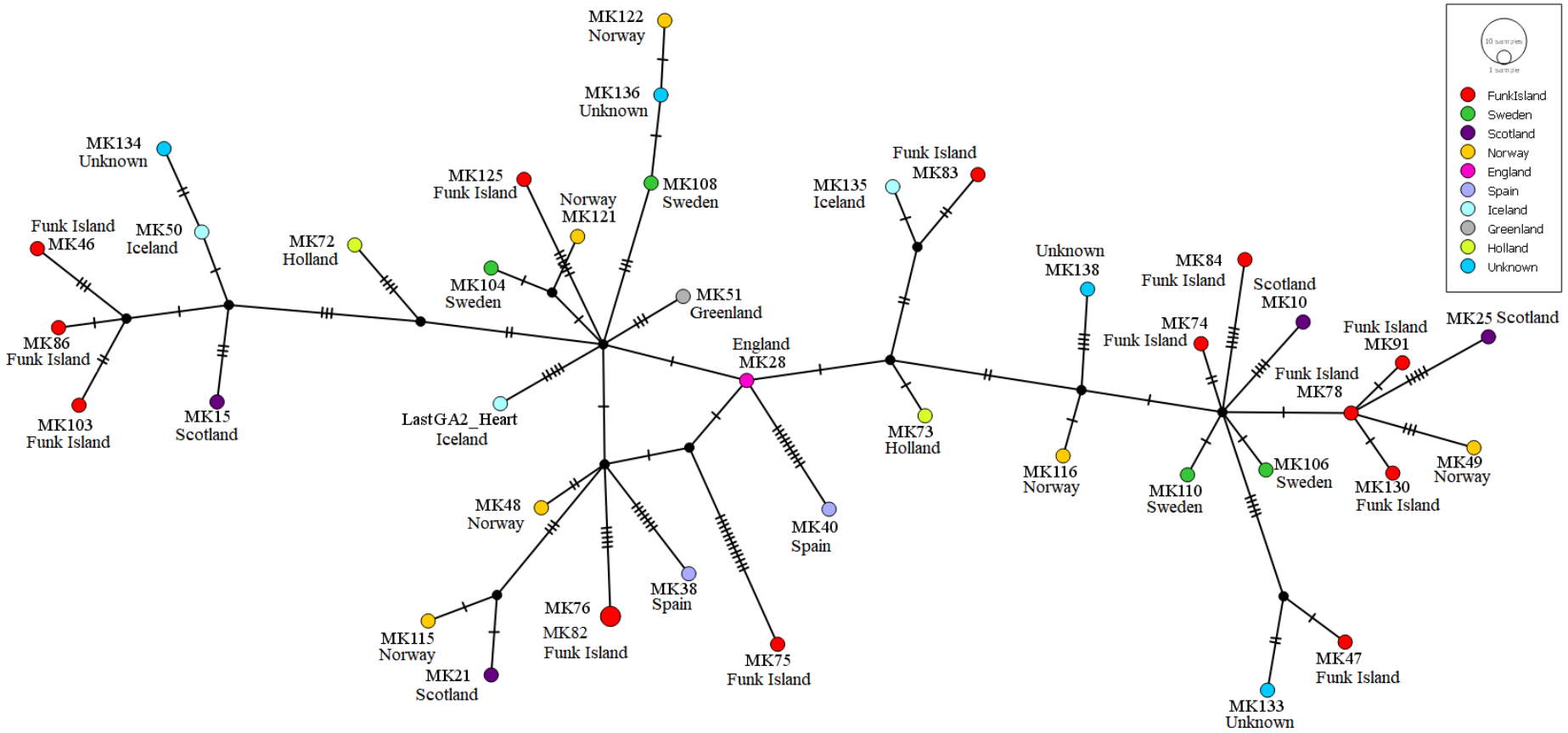

Figure 3.4 A Median Joining Network (Bandelt *et al*., 1999) created in PopART (Leigh & Bryant, 2015), using the mitogenome data from all great auks used the in analysis.

3.3.3 Demographic inferences using unpartitioned data set

Bayes factor (BF) model comparisons showed that the unpartitioned data was best described by a relaxed lognormal molecular clock (log10 BF: 2.71 compared to strict clock) model. To infer the best supported coalescent prior, we compared the flexible Bayesian Skyline Plot model, which allows population size to vary over time, an exponential growth tree prior and a simpler constant population size model. The BF support for a Bayesian skyline plot indicating a population size increase over the past 250,000 years was "substantial" to "strong" (log10 BF: 0.98) but not "decisive" (Kass and Raftery, 1995). Likewise, for the exponential growth tree prior the BF support was "strong" (log10 BF: 1.122) but not "decisive". Therefore, the BSP and exponential growth prior were rejected in favour of the simpler constant population size coalescent prior.

While our main calibration information came from the substitution rates calculated by Nabholz et al. (2016), however, our samples were not all of the same age. To test if this had an impact on our results, we ran the analysis including and excluding the age of individual samples (tip dates). When viewed in Tracer v1.5.0, treeModel.rootHeight was 0.543, with tip dates and 0.541, without tip dates. This showed that calibrations did not change the demographic inference or estimated divergence times. All further analyses were conducted with tip calibrations as well as a uniform prior of 0.002-0.003 subst/site/Myrs. The mean effective female population size (N_e) was estimated at 310,000 (HPD 174,000- 467,000).

3.3.4 Partitioned dataset

Our Bayesian phylogenetic analysis of the dated mitogenomes produced a posterior median estimate of 42,188 years (95% credibility interval 24,743–84,894 years) for the age of the most recent common ancestor. The product of the effective female population size and generation time was estimated at 87,973 (95% credibility interval 29,724–233,900). The median posterior estimate of the mutation rate was 2.74×10^{-8} mutations/site/year (95% credibility interval 9.83×10^{-9}–4.53×10^{-8}). The data set showed some evidence of temporal structure, passing the more lenient criterion CR1 but not the more stringent CR2 of the date-randomization test (Supplementary Materials Fig. S3.2; (Duchêne *et al.*, 2015)).

When the undated mitogenomes were included, with constraints on their ages informed by their sampling contexts, the Bayesian rate estimate was 2.86×10^{-8} mutations/site/year (95%

credibility interval 1.34×10^{-8}–4.27×10^{-8}). The estimates from our Bayesian phylogenetic approach were lower than that obtained by regression of root-to-tip distances against sampling times, which produced a point estimate of 6.30×10^{-8} mutations/site/year. However, least-squares dating produced a lower rate estimate of 9.50×10^{-9} mutations/site/year.

3.3.5 Extinction simulation using population viability analysis software

The results of the PVA simulations showed that regardless of extinction definition, egg harvest size or total population size, a harvest rate of 2% generates a 0% probability of extinction and at harvest rates higher than 6% extinction occurs in 100% of simulation replications within a short period of time (Table 3.2, 3.3, 3.4). Simulations with a harvest rate of 5% of the population (i.e. 50,000 birds from a population of 1million, 150,000 birds for a population size of 3 million) cross a threshold and begin to have an observable effect on the probability of extinction. Simulations were unaffected by extinction definition and therefore all results reported are from simulations run under extinction defined as 'only 1 sex remains'.

For a population size of one million, with a harvest rate of 5% and 50,000 eggs harvested, we see a significant number of simulation replications going extinct (77% of 100 simulations, mean time to extinction 162 years). When we looked at the impact of harvesting 3% and 4% of the estimated population size of one million birds, we see that with 3% and 50,000 eggs there were zero simulations going extinct but when 100,000 eggs were harvested, it rose to 34% of simulation replications going extinct in 178 years. With a 4% harvest rate with 100,000 eggs removed, 99% of simulation replications went extinct when the population size was one million. With a 5% harvest rate, we see a 77% or 100% chance of extinction for our population size of 1 million with harvesting 50,000 or 100,000 eggs, respectively. For our simulations with initial population size of three million and ten million, we see a significant number of the simulation replications going extinct once the harvest rate reaches 6%. With regards to the impact that the two levels of egg harvesting have on the probability of extinction, harvesting 100,000 eggs does increase the probability of extinction and decreases the time to extinction, especially in the smaller population size of one million birds, but the same conclusions regarding percentage of the population harvested are still applicable i.e. for a population size of three and ten million at 6% harvesting, a significant number of simulations will end in extinction.

Population size	Harvest rate %	Number of birds	Number of eggs	Probability of extinction	Mean time to extinction	Median time to extinction
1,000,000	2	20,000	50,000	0.00	0.00	0.00
1,000,000	3	30,000	50,000	0.00	0.00	0.00
1,000,000	4	40,000	50,000	0.08	169.50	-
1,000,000	5	50,000	50,000	0.77	161.91	184.00
1,000,000	6	60,000	50,000	1.00	51.99	41.00
1,000,000	7	70,000	50,000	1.00	27.64	25.00
1,000,000	10	100,000	50,000	1.00	14.06	14.00
1,000,000	2	20,000	100,000	0.00	0.00	0.00
1,000,000	3	30,000	100,000	0.34	178.03	-
1,000,000	4	40,000	100,000	0.99	107.15	87.00
1,000,000	5	50,000	100,000	1.00	39.26	32.00

Table 3.2 Results of hunting simulations for using a population size of 1 million birds, where extinction was defined as 'only 1 sex remains'. Harvest rate describes the percent of population size harvested annually. Number of birds is the total number of birds killed annually, which was split between the age cohorts as described in the methods, number of eggs is the number of eggs harvested annually (50,000 reported for all harvest rates and 100,000 for 2%-5%). Mean and median time to extinction are reported in years.

Population size	Harvest rate %	Number of birds	Number of eggs	Probability of extinction	Mean time to extinction	Median time to extinction
3,000,000	2	60,000	50,000	0.00	0.00	0.00
3,000,000	5	150,000	50,000	0.07	114.86	-
3,000,000	6	180,000	50,000	0.80	169.03	201.00
3,000,000	7	210,000	50,000	1.00	49.74	43.00
3,000,000	10	300,000	50,000	1.00	17.75	17.00
3,000,000	5	150,000	100,000	0.31	182.71	-
3,000,000	6	180,000	100,000	0.99	77.10	70.00
3,000,000	7	210,000	100,000	1.00	37.63	32.00

Table 3.3 Results of hunting simulations for using a population size of 3 million birds, where extinction was defined as 'only 1 sex remains'. Harvest rate describes the percent of population size harvested annually. Number of birds is the total number of birds killed annually, which was split between the age cohorts as described in the methods, number of eggs is the number of eggs harvested annually (50,000 reported for all harvest rates and 100,000 for 5%, 6% and 7%). Mean and median time to extinction are reported in years.

Population size	Harvest rate %	Number of birds	Number of eggs	Probability of extinction	Mean time to extinction	Median time to extinction
10,000,000	2	200,000	50,000	0.00	0.00	0.00
10,000,000	5	500,000	50,000	0.01	123.00	-
10,000,000	6	600,000	50,000	0.42	165.79	0.00
10,000,000	7	700,000	50,000	1.00	65.56	57.00
10,000,000	10	1,000,000	50,000	1.00	18.77	18.00
10,000,000	5	500,000	100,000	0.02	112.00	-
10,000,000	6	600,000	100,000	0.69	160.52	215.00
10,000,000	7	700,000	100,000	1.00	51.72	42.00

Table 3.4 Results of hunting simulations for using a population size of 10 million birds. Harvest rate describes the percent of population size harvested annually. Number of birds is the total number of birds killed annually, number of eggs is the number of eggs harvested annually (50,000 reported for all harvest rates and 100,000 for 5%, 6% and 7%). Mean and median time to extinction are reported in years.

3.3.6 GPS-equipped drifting capsules

Data collected from the GPS-equipped drifting capsules shows the routes taken around the North Atlantic Ocean (Fig. 3.5).

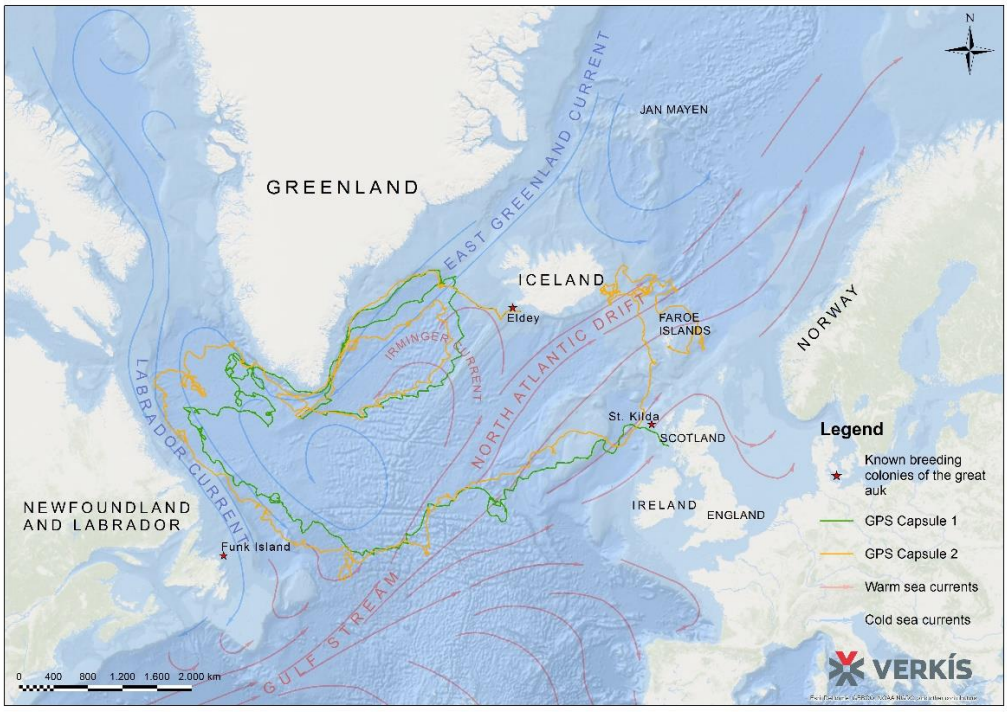

Figure 3.5 GPS data from two capsules (green and yellow lines) deployed in the North Atlantic, showing possible routes that the great auk may have used to move between colonies, aided by the ocean currents, waves and wind.

Forces driving the capsules are currents, wind and waves. The route taken demonstrates how great auks could have used these forces to aid movements around the North Atlantic, between colonies, during the months they were at sea. A full description of the route taken can be found in Supplementary Materials S3.4.

3.4 Discussion

3.4.1 Cause of extinction

Our findings provide strong evidence that the great auk was a species that was not in decline or necessarily at risk of extinction prior to the intense human exploitation. The information gained from examining past genetic structure and demography indicates that great auk population sizes remained constant through time, as shown from analysis with and without data partitioning. The effective female population size from the full unpartitioned dataset was high at ~310,000, indicating that the census size could have been in the millions, as suggested by Birkhead (1993). Based on our analyses, there was no detectable decline in the N_e that would correspond to any documented climatic events, such as the 'Little Ice Age', and we observed high genetic diversity across our sample range, right up to the species demise. However, it has been well documented that on the eastern side of the North Atlantic a decline in great auk numbers did occur over time. This is evident from the decline in bone finds from about 0CE in England, Scotland and Scandinavia (Hufthammer, 1982; Bengtson, 1984; Serjeantson, 2001). There was also no evidence that the species' experienced serial bottlenecks across a representative geologically recent time period. Our genetic data failed to detect any population structure in space or time (Fig. 3.3, 3.4), indicating no barrier to breeding between populations across the range of the species. This result is in contrast to what we may have expected to see based on the suggestions of limited or no interbreeding between populations of either side of the North Atlantic, oceanographic-related size differences (Burness and Montevecchi, 1992), and the suspected regional philopatry in this species (Bengtson, 1984; Montevecchi and Kirk, 1996). Such a lack of structure is, however, common in sea birds, and has been observed in several great auk relatives e.g. thick-billed murre (*Uria lomvia*) (no structure within ocean basins) (Tigano *et al.*, 2015), common murre (*Uria aalge*) (structure in Atlantic but not in Pacific) (Morris-Pocock *et al.*, 2008), ancient murrelets (*Synthliboramphus antiquus*) (no genetic differentiation in North Pacific) (Pearce *et al.*, 2009), and little auk (*Alle alle*) (no structure in Arctic) (Wojczulanis-Jakubas *et al.*, 2014). Findings are further supported by our data collected from GPS-enabled drifting capsules released into the North Atlantic. The data

shows that it would have been possible for great auks to move very easily between the various known breeding grounds (e.g. Funk Island, St Kilda and Eldey Island (Fig. 3.5)) by using the oceanic currents to aid migration, as suggested by Brown (1985) (Meldgaard, 1988). High mobility is further supported by the great auks' ability to track its habitat in response to climate change. The archaeological record shows that the great auk distribution changed over time, with evidence of both northward and southward movements (Bengtson, 1984; Meldgaard, 1988; Serjeantson, 2001; Campmas *et al.*, 2010). Southern movements are evident from bones, some of which show cut marks indicative of hunting, and even cave paintings, that show great auks were found in southern Italy, Portugal, Spain and France, during the Upper Pleistocene and Holocene and even Morocco (Holocene) (Serjeantson, 2001; Campmas *et al.*, 2010). On the North American side of the Atlantic, evidence of the great auk on the eastern seaboard can be found as far south as Florida (Hay, 1902; Serjeantson, 2001). There is evidence to suggest that during warmer periods of the 12^{th}-13^{th} century there was movement of breeding birds northwards into Greenland (Serjeantson, 2001). Such an ability to track habitat has been reported for a number of species and is commonly considered an advantageous trait that buffers the effects of environmental change (e.g. moa (Rawlence *et al.*, 2012), bowhead whale (Foote *et al.*, 2013), sea lions and penguin species (Waters *et al.*, 2017), ptarmigan species (Lagerholm *et al.*, 2017) cf. arctic fox (Dalén *et al.*, 2007)).

Collectively our findings suggest that the great auk was not a species that was at risk of extinction at the time when human hunting intensified. Despite possessing life history traits that are said to 'promote' extinction risk (see McKinney, 1997) or perhaps render it over-specialised and unfit for survival, (such as flightlessness, low fecundity, and restrictions with regard to suitable breeding grounds), our results suggest the great auk was well adapted for surviving, even through periods of environmental change. Similar to the extinction of New Zealand moa species (Allentoft *et al.*, 2014; Perry *et al.*, 2014), it appears that human hunting was a primary driving factor in the demise of this iconic North Atlantic sea bird. Given the reported and estimated large population size and the extensive geographical range of the species, the main question therefore becomes how intense the hunting pressure was to bring about extinction.

3.4.2 How intense was the hunting?

Our results of the PVA analysis indicate that for a population size of one million birds, a number which is not unrealistic as justified in the methods (Section 3.2.9), harvesting just 40,000 or 50,000 birds annually (depending on egg harvesting rate) would cause certain extinction. Even for a population size of three million birds, which again is a plausible estimate, the killing of 180,000 birds a year, led to extinction in almost all our simulations within significantly less than 350 years. Similarly, for a conservatively large census size of ten million birds, a harvesting rate of 7% (i.e. 700,000 birds) would lead to certain extinction. While at first such numbers appear high, they are not unreasonable based on the contemporary accounts of the slaughtering on Funk Island. The archaeological record and numerous historical accounts are testimony to the size of great auk populations and the scale of hunting. From 1497CE when the Europeans discovered the rich fishing grounds of Newfoundland, year on year fleets of 300-400 ships from various European countries were drawn to the region (Steenstrup, 1855; Bengtson, 1984). Fishing stations were set up near seabird colonies and the colonies were heavily exploited (Pope, 2009). Numerous historic reports summarised by Gaskell (2000) describe filling boats with several tons of birds in less than half an hour, in addition to the barrels of eggs that were also taken. As well as those killed on land, there are reports of great auks and other seabirds being caught by fishing lines while at sea (Pope, 2009). Similarly, great auks were likely caught in fishing nets as is seen in many extant diving bird species (Piatt & Nettleship, 1985, 1987; Montevecchi & Kirk, 1996). Later, great auks were specifically targeted for the feather trade. Hunters would live on Funk Island all summer with the purpose of killing the birds (Kirkham and Montevecchi, 1982; Gaskell, 2000). To add to the effects of excessive hunting, the great auk laid only one egg a year, which was not replaced if removed. Thus, replenishing the large number of auks lost annually would have been highly improbable (Gaskell, 2000).

If we compare our harvest estimates of great auks to those of extant species today, we again see that our results do not provide an unrealistic picture. For example, off the coast of Newfoundland and Canada, between 200,000-300,000 murres (*Uria spp*.) are killed legally every year. This number was even higher before the mid-1990s, when between 300,000-700,000 thick-billed murres alone were being harvested annually (Wilhelm *et al.*, 2008). In Iceland, the numbers of Atlantic puffins killed annually, are 150,000-233,000, which is approximately 2-3% population, but for species of black-backed gulls between 25-30% of the

population are killed annually (Merkel and Barry, 2008). While the current figures for auk species harvested annually are now significantly less than this and continue to decline, those above show that the hunting pressure required to result in the extinction of a population of one, three or ten million great auks, would not be considered excessive even by modern standards.

3.5 Conclusion

The role of humans versus environmental change in species extinction has been long debated, not only for the great auk, but also across other documented species losses (Shapiro *et al.*, 2004; Lorenzen *et al.*, 2011; Cooper *et al.*, 2015). Additionally, it is recognised that obtaining a better understanding of causes of previous extinction is applicable to contemporary conservation measures. The majority of aDNA studies that investigate cause of extinction focus on megafaunal species (e.g. mammoth) that became extinct thousands of years ago. However, the great auk is one of very few seabirds that went extinct in the recent past. Therefore, understanding its extinction has greater implications for our understanding of present-day threats to biodiversity, especially to its extant relatives. In contrast to most studies on Pleistocene extinctions, which have argued for at least some level of climate-driven environmental contributions to species extinction, we could not find any evidence of factors other than human hunting that led to the extinction of the great auk. The great auk therefore shows that industrial scale commercial exploitation of natural resources can drive even an abundant, wide ranging, highly mobile and genetically diverse species, to extinction within short periods of time. It emphasises the need for thorough monitoring of commercial harvesting of species, in particular in poorly researched environments such as our oceans, to create sustainable ecosystems for the future and ensure evidence based conservation management of biodiversity.

3.6 Additional information

This paper is currently in preparation for publication.

3.6.1 Tentative author list

Jessica E. Thomas, Gary R. Carvalho[†], James Haile, Nicolas J. Rawlence, Michael D. Martin, Simon Y. W. Ho, Arnór Þ. Sigfússon, Vigfús A. Jósefsson, Morten Frederiksen, Jannie F. Linnebjerg, Jose A. Samaniego Castruita, Jonas Niemann, Mikkel-Holger S. Sinding, Marcela

Sandoval-Velasco, André E. R. Soares, Christina Barilaro, Julia Best, Dirk Brandis, Chiara Cavallo, Mikelo Elorza, Kimball L. Garrett, Maaike Groot, Friederike Johansson, Jan T. Lifjeld, Göran Nilson, Dale Serjeanston, Paul Sweet, Errol Fuller, Anne Karin Hufthammer, Morten Meldgaard, Jon Fjeldså, Beth Shapiro, Michael Hofreiter, John R. Stewart[†], M. Thomas P. Gilbert[†] and Michael Knapp[†] ([†]Equal contribution)

3.6.2 Author contributions

J.E.T., G.R.C., N.J.R., M.H., J.R.S., M.T.P.G., and M.K. conceived the study, J.E.T., J.H., G.R.C., M.T.P.G., and M.K. designed the experiments; J.E.T., J.H., M-H.S.S., and M.S.-V. conducted the experiments; J.H. and J.A.S.C. designed the bait; A.Þ.S., V.A.J., M.F., and J.F.L. provided GPS buoy data and associated text; J.E.T., M.D.M., S.Y.W.H, A.E.R.S., J.A.S.C., and J.N. analysed the data; C.B., J.B., D.B., C.C., M.E., K.L.G., M.G., F.J., J.T.L., G.N., D.S., P.S., A.K.H., provided samples used in the study; N.J.R., M.H., and J.R.S. provided the initial framework for the study; all authors contributed to ideas and writing the paper.

Chapter 3: Supplementary Materials

S3 Tables

Lab ID	Institution	Institution Number (if available)	Site Location & Country	Age (associated period / estimated median age)
MK09	Bournemouth University	Spec No 22722. CH98 (595)	Cladh Hallan, South Uist, Scotland	1135-1035 BC-635-535 BC est. 2850
MK10	Bournemouth University	CH99 (692)	Cladh Hallan, South Uist, Scotland	1135-1035 BC-635-535 BC est. 2850
MK11	Bournemouth University	CH00 (1512)	Cladh Hallan, South Uist, Scotland	1135-1035 BC-635-535 BC est. 2850
MK12	Bournemouth University	CH96 (423)	Cladh Hallan, South Uist, Scotland	1135-1035 BC-635-535 BC est. 2850
MK13	Bournemouth University	Spec No 20232 CH00 (1785)	Cladh Hallan, South Uist, Scotland	1135-1035 BC-635-535 BC est. 2850
MK14	Bournemouth University	(36) B96 337	Bornais Mound 1, South Uist, Scotland	Late Scottish Iron Age 400-800AD est. 1500
MK15	Bournemouth University	(481) 490 2	Bornais Mound 1, South Uist, Scotland	Late Scottish Iron Age 400-800AD est. 1500
MK16	Bournemouth University	(90) (404) B97	Bornais Mound 1, South Uist, Scotland	Late Scottish Iron Age 400-800AD est. 1500
MK17	Bournemouth University	(487) 45 456	Bornais Mound 1, South Uist, Scotland	Late Scottish Iron Age 400-800AD est. 1500
MK18	Bournemouth University	(481) 307	Bornais Mound 1, South Uist, Scotland	Late Scottish Iron Age 400-800AD est. 1500
MK19	University of Southampton	255 (Context 027)	Tofts Ness, Sanday, Orkney, Scotland	Neolithic/Early Bronze Age (2000BC) est. 4000
MK20	University of Southampton	3 (TN86, 016)	Tofts Ness, Sanday, Orkney, Scotland	Neolithic/Early Bronze Age (2000BC) est. 4000
MK21	University of Southampton	57 (Context 1123)	Tofts Ness, Sanday, Orkney, Scotland	Neolithic/Early Bronze Age (2000BC) est. 4000
MK22	University of Southampton	366 (Context 1007)	Tofts Ness, Sanday, Orkney, Scotland	Neolithic/Early Bronze Age (2000BC) est. 4000
MK23	University of Southampton	136 (TN86, 012)	Tofts Ness, Sanday, Orkney, Scotland	Neolithic/Early Bronze Age (2000BC) est. 4000
MK24	University of Southampton	293 (TN 0194)	Tofts Ness, Sanday, Orkney, Scotland	Neolithic/Early Bronze Age (2000BC) est. 4000
MK25	University of Southampton	108 (TN86, 0177)	Tofts Ness, Sanday, Orkney, Scotland	Neolithic/Early Bronze Age (2000BC) est. 4000
MK26	Portland Museum	PM1 (Context 189)	Royal Manor Field, Portland, England	Roman age est. 2000
MK27	Portland Museum	PM2 (Context 189)	Royal Manor Field, Portland, England	Roman age est. 2000
MK28	Portland Museum	PM3 (Context 199)	Royal Manor Field, Portland, England	Roman age est. 2000

Lab ID	Institution	Institution Number (if available)	Site Location & Country	Age (associated period / estimated median age)
MK29	Portland Museum	PM4 (Context 199)	Royal Manor Field, Portland, England	Roman age est. 2000
MK30	National Museum of Scotland	NA	Middle Oronsay, Scotland	6200-5100
MK31	National Museum of Scotland	1940-15-40	Knowe of Ramsay, Orkney, Scotland	Neolithic
MK32	National Museum of Scotland	z.2002.175.1 GA753	Freswick Links, Caithness, Scotland	Late Iron Age
MK36	Arkeologi Museoa, Bilbao	SC. B4.73.14	Santa Catalina, Lekeitio, Spain	Magda. Sup. Level II 12500-14000 est. 13250
MK37	Arkeologi Museoa, Bilbao	SC.C8. L 22.295	Santa Catalina, Lekeitio, Spain	Magda. Sup. Level II 12500-14000 est. 13250
MK38	Arkeologi Museoa, Bilbao	SC.B6.L26.51	Santa Catalina, Lekeitio, Spain	Magda. Sup. Level II 12500-14000 est. 13250
MK39	Arkeologi Museoa, Bilbao	SC.C6.L29E.1602	Santa Catalina, Lekeitio, Spain	Magda. Fin. Level III 12900-15000 est. 13950
MK40	Arkeologi Museoa, Bilbao	SC.C6.L31C.1428	Santa Catalina, Lekeitio, Spain	Magda. Fin. Level III 12900-15000 est. 13950
MK41	Arkeologi Museoa, Bilbao	SC.B8.L29E.2601	Santa Catalina, Lekeitio, Spain	Magda. Fin. Level III 12900-15000 est. 13950
MK42	Arkeologi Museoa, Bilbao	SC.C8.L29G.10440-10442	Santa Catalina, Lekeitio, Spain	Magda. Fin. Level III 12900-15000 est. 13950
MK43	Arkeologi Museoa, Bilbao	SC.B6.L29E.1754	Santa Catalina, Lekeitio, Spain	Magda. Fin. Level III 12900-15000 est. 13950
MK44	Arkeologi Museoa, Bilbao	SC.B8.147.425	Santa Catalina, Lekeitio, Spain	Magda. Fin. Level III 12900-15000 est. 13950
MK45	Arkeologi Museoa, Bilbao	PRA.B8.7365	Pico Ramos, Muskiz, Spain	5,000
MK46	Natural History Museum of Denmark	NA	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK47	Natural History Museum of Denmark	5504	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK48	Natural History Museum of Denmark	10010	Vardø, Norway	Unknown, but unlikely over 5000
MK49	Natural History Museum of Denmark	NA	Vardø, Norway	Unknown, but unlikely over 5000
MK50	Natural History Museum of Denmark	NA	Site location unknown, Iceland	Killed 1821
MK51	Natural History Museum of Denmark	NA	Qeqertarsuatsiaat, Greenland	Killed 1815
MK52	Natural History Museum of Denmark	NA	Funk Island, Newfoundland, Canada	Unknown, unlikely over 1000
MK53	Natural History Museum of Denmark	ZMK 4/1861 (19/63)	Mejlgård, Randers, Denmark	Ertebølle culture. 5300 BC – 3950 BC C-14: (*Bos primigenius*) 5115 ± 70 BP
MK54	Natural History Museum of Denmark	ZMK 4/1893 (24/63)	Havnø, Aalborg, Denmark	Ertebølle culture. C-14: 4130 ± 40 BP

Lab ID	Institution	Institution Number (if available)	Site Location & Country	Age (associated period / estimated median age)
MK55	Natural History Museum of Denmark	ZMK 130/1967	Villingebæk Island, Frederiksborg, Denmark	Kongemose Culture, 6000 BC–5200 BC
MK56	Natural History Museum of Denmark	ZMK 101/1951	Reykjavik, Iceland	Unknown
MK57	Natural History Museum of Denmark	ZMK 11/1915	Nivågård, Frederiksborg, Denmark	Early Ertebølleruten Culture
MK58	Natural History Museum of Denmark	ZMK 111/1952	Kangeq, Greenland	Thule Culture 1400-1952 AD
MK59	Natural History Museum of Denmark	ZMK 111/1952	Kangeq, Greenland	Thule Culture 1400-1952 AD
MK60	Natural History Museum of Denmark	ZMK 1/1851	Gudumlund, Aalborg, Denmark	Ertebølle culture. C-14: *Phoca groenl* 5275 ± 50 BP
MK61	Natural History Museum of Denmark	ZMK 1/1851	Gudumlund, Aalborg, Denmark	Ertebølle culture. C-14: *Phoca groenl* 5275 ± 50 BP
MK62	Natural History Museum of Denmark	ZMK 31/1901	Sølager, Frederiksborg, Denmark	Ertebølleruten Culture, Neolithic Funnel Beaker Culture. C-14: *Phoca groenl*. 5460 ± 40 BP
MK63	Natural History Museum of Denmark	ZMK 08/1908	Kassemose, Frederiksborg, Denmark	Ertebølleruten Culture and Neolithic
MK64	Natural History Museum of Denmark	ZMK 104/1968	Olsbjerg, Sejerø, Holbæk, Denmark	Assumed Ertebølleruten Culture
MK65	Natural History Museum of Denmark	ZMK 5/1898	Klintesø, Holbæk, Denmark	Ertebølleruten Culture C-14: Capra: 3926 ± 33 BP
MK66	Natural History Museum of Denmark	ZMK 103/1956	Bergmandsdal, Frederiksborg, Denmark	Ertebølleruten Culture
MK67	Natural History Museum of Denmark	ZMK 130/1970	Hope Colony, Godthåbsfjord, Greenland	1721-1728AD 296-289
MK68	Natural History Museum of Denmark	ZMK 130/1970	Hope Colony, Godthåbsfjord, Greenland	1721-1728AD 296-289
MK69	Natural History Museum of Denmark	ZMK 39/1909	Kollafjarðarnes, Iceland	Unknown
MK70	Natural History Museum of Denmark	ZMK 39/1909	Vardø, Norway	1475 - 1700 century
MK71	Natural History Museum of Denmark	ZMK 142/1972	Igdlorpait, Greenland	1721-1728AD 296-289
MK72	University of Amsterdam	0B97-2159-1	Schipluiden, Midden-Delfland, Netherlands	First century AD (Roman) 1600-2000
MK73	University of Amsterdam	NA	Velsen, Netherlands	First century AD (Roman) 1600-2000
MK74	American Museum of Natural History	Bone # 4547, AMNH 31855	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK75	American Museum of Natural History	Bone # 4408, AMNH 31847	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK76	American Museum of Natural History	Bone # 4218, AMNH 31833	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000

Lab ID	Institution	Institution Number (if available)	Site Location & Country	Age (associated period / estimated median age)
MK77	American Museum of Natural History	Bone # 4387, AMNH 31846	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK78	American Museum of Natural History	Bone # 4624, AMNH 31857	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK79	American Museum of Natural History	Bone # 4339, AMNH 31843	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK80	American Museum of Natural History	Bone # 4549, AMNH 31856	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK81	American Museum of Natural History	Bone # 4446, AMNH 31851	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK82	American Museum of Natural History	Bone # 4204, AMNH 31830	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK83	American Museum of Natural History	Bone # 4285, AMNH 31842	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK84	American Museum of Natural History	Bone # 4179, AMNH 31829	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK85	American Museum of Natural History	Bone # 4170, AMNH 31828	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK86	American Museum of Natural History	Bone # 4269, AMNH 31841	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK87	American Museum of Natural History	Bone # 4230, AMNH 31836	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK88	American Museum of Natural History	Bone # 4448, AMNH 31852	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK89	American Museum of Natural History	Bone # 4534, AMNH 31854	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK90	American Museum of Natural History	Bone # 4526, AMNH 31853	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK91	American Museum of Natural History	Bone # 4423, AMNH 31848	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK92	American Museum of Natural History	Bone # 4425, AMNH 31849	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK93	American Museum of Natural History	Bone # 4429, AMNH 31850	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK94	American Museum of Natural History	Bone # 4211, AMNH 31831	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK95	American Museum of Natural History	Bone # 4214, AMNH 31832	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK96	American Museum of Natural History	Bone # 4219, AMNH 31834	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK97	American Museum of Natural History	Bone # 4223, AMNH 31835	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK98	American Museum of Natural History	Bone # 4232, AMNH 31837	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK99	American Museum of Natural History	Bone # 4240, AMNH 31838	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK100	American Museum of Natural History	Bone # 4254, AMNH 31839	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000

Lab ID	Institution	Institution Number (if available)	Site Location & Country	Age (associated period / estimated median age)
MK101	American Museum of Natural History	Bone # 4264, AMNH 31840	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK102	American Museum of Natural History	Bone # 4374, AMNH 31844	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK103	American Museum of Natural History	Bone # 4380, AMNH 31845	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK104	Gothenburg Naturhistoriska Museum	GNM 6022	Sotenkanalen, Sweden	est. 5000
MK105	Gothenburg Naturhistoriska Museum	GMN 6048	Sotenkanalen, Sweden	est. 5000
MK106	Gothenburg Naturhistoriska Museum	GNM 18684	Nödö, Sweden	est. 5000
MK107	Gothenburg Naturhistoriska Museum	GNM 6112	Nödö, Sweden	est. 5000
MK108	Gothenburg Naturhistoriska Museum	GNM 6849	Skalbank Otterön, Sweden	est. 5000
MK109	Gothenburg Naturhistoriska Museum	GNM 6194	Sotenkanalen, Sweden	est. 5000
MK110	Gothenburg Naturhistoriska Museum	GNM 18,685	Skalbank Otterön, Sweden	est. 5000
MK111	Gothenburg Naturhistoriska Museum	GNM 9644	Skalbank Otterön, Sweden	est. 5000
MK112	Gothenburg Naturhistoriska Museum	84,152-16 Hultas, A 6	Rottjärnslid, Sweden	est. 5000
MK113	Gothenburg Naturhistoriska Museum	84,152-16 Hog 5	Rottjärnslid, Sweden	est. 5000
MK114	Gothenburg Naturhistoriska Museum	GNM 18,692	Sotenkanalen, Sweden	est. 5000
MK115	University of Bergen	JS.603	Iversfjord, Finnmark, Norway	3000-2400BC, est. 5016-4416
MK116	University of Bergen	JS.523	Storbåthelleren, Nordland, Norway	3300-2790/790-100BC, est. 5316-2116
MK117	University of Bergen	JS.528	Storbåthelleren, Nordland, Norway	3300-2790/790-100BC, est. 5316-2116
MK118	University of Bergen	JS.412	Viste, Rogaland, Norway	Mesolithic 7000-5000BC
MK119	University of Bergen	JS.412 3	Viste, Rogaland, Norway	Mesolithic 7000-5000BC
MK120	University of Bergen	JS.427	Skjonghelleren, Møre og Romsdal, Norway	Older Iron Age 500BC-500AD
MK121	University of Bergen	JS.371	Nyelv, Finnmark, Norway	5000-4000BP C-14: 4160+- 80
MK122	University of Bergen	JS.260	Kirkehelleren, Nordland, Norway	Mesolithic to 500AD
MK123	University of Bergen	JS.445	Gronehelleren, Sogn og Fjordane, Norway	Late Mesolithic- recent times

Lab ID	Institution	Institution Number (if available)	Site Location & Country	Age (associated period / estimated median age)
MK124	University of Bergen	JS.445	Gronehelleren, Sogn og Fjordane, Norway	Late Mesolithic- recent times
MK125	Zoological Museum Oslo	1213 (1)	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK126	Zoological Museum Oslo	1213 (2)	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK127	Zoological Museum Oslo	1213 (3)	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK128	Zoological Museum Oslo	1213 (4)	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK129	Zoological Museum Oslo	1213 (5)	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK130	Zoological Museum Oslo	1213 (6)	Funk Island, Newfoundland, Canada	Unknown, but unlikely over 1000
MK131	Natural History Museum of Denmark	NHMD 153069	Eldey Island, Iceland	Killed 1844
MK132	Natural History Museum of Denmark	NHMD 153070	Eldey Island, Iceland	Killed 1844
MK133	Landesmuseum Natur und Mensch Oldenburg	AVE 8086	Unknown, most likely Iceland	est. 170-200 years old
MK134	Übersee-Museum Bremen	RKNr. 2392	Unknown, most likely Iceland	est. 170-200 years old
MK135	Institut royal des Sciences naturelles de Belgique	RBINS 5355	Eldey Island, Iceland	Killed 1844
MK136	Natural History Museum of Los Angeles County	LACM 76476	Unknown, most likely Iceland	est. 170-200 years old
MK137	Musée d'Histoire Naturelle Neuchâtel	NA	Unknown, most likely Iceland	est. 170-200 years old
MK138	Zoologisches Museum der Christian-Albrechts Universität zu Kiel	cat. No. A0585	Unknown, most likely Iceland	est. 170-200 years old
LastGA2_ Heart	Natural History Museum of Denmark	NHMD 153070	Eldey Island, Iceland	Killed 1844

Table S3.1 Sample information for samples used in the mitochondrial genome study. Information listed shows institution name and number where sample was sourced, the site location and country where sample was discovered (if known) and any associated date/age information known.

Lab ID	Sample type	Weight (mg)	Extraction	DSlib	SSlib	Capture	Sequencing type
MK09	Bone	33.6	Yes	Yes	Yes	Yes	NA
MK10	Bone	20.1	Yes	Yes	Yes	Yes	MiSeq PE75
MK11	Bone	27.9	Yes	Yes	No	No	NA
MK12	Bone	32.8	Yes	Yes	No	No	NA
MK13	Bone	7.6	Yes	Yes	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK14	Bone	21.3	Yes	Yes	No	No	NA
MK15	Bone	31.0	Yes	Yes	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK16	Bone	26.2	Yes	Yes	No	No	NA
MK17	Bone	48.6	Yes	Yes	No	No	NA
MK18	Bone	22.1	Yes	Yes	Yes	Yes	NA
MK19	Bone	15.3	Yes	Yes	No	No	NA
MK20	Bone	27.8	Yes	Yes	No	No	NA
MK21	Bone	22.2	Yes	Yes	Yes	Yes	MiSeqPE75
MK22	Bone	19.6	Yes	Yes	No	No	NA
MK23	Bone	35.0	Yes	Yes	No	No	NA
MK24	Bone	53.6	Yes	Yes	No	No	NA
MK25	Bone	38.2	Yes	Yes	Yes	Yes	MiSeq PE150/ HiSeq SR100
MK26	Bone	50.0	Yes	Yes	Yes	Yes	HiSeq SR100
MK27	Bone	69.1	Yes	Yes	Yes	Yes	NA
MK28	Bone	34.8	Yes	Yes	Yes	Yes	MiSeq PE150/ HiSeq SR100
MK29	Bone	59.6	Yes	Yes	Yes	Yes	HiSeq SR100
MK30	Bone	36.5	Yes	Yes	Yes	Yes	Nano MiSeq PE70/ MiSeq PE 75
MK31	Bone	32.4	Yes	Yes	Yes	Yes	Nano MiSeq PE70/ MiSeq PE 75
MK32	Bone	63.2	Yes	Yes	Yes	Yes	NA
MK33	Bone	31.0	Yes	Yes	Yes	Yes	NA
MK34	Bone	41.7	Yes	Yes	Yes	Yes	Nano MiSeq PE70/ MiSeq PE 75
MK35	Bone	47.7	Yes	Yes	No	No	NA
MK36	Bone	36.2	Yes	Yes	Yes	Yes	NA
MK37	Bone	67.3	Yes	Yes	Yes	Yes	HiSeq SR100
MK38	Bone	58.0	Yes	Yes	Yes	Yes	HiSeq SR100
MK39	Bone	46.4	Yes	Yes	No	No	NA
MK40	Bone	59.0	Yes	Yes	Yes	Yes	MiSeq PE150/ HiSeq SR100
MK41	Bone	56.2	Yes	Yes	Yes	Yes	No
MK42	Bone	40.2	Yes	Yes	Yes	Yes	MiSeq PE75
MK43	Bone	55.9	Yes	Yes	No	No	NA
MK44	Bone	58.6	Yes	Yes	Yes	Yes	Nano MiSeq PE70/ MiSeq PE 75
MK45	Bone	89.2	Yes	Yes	Yes	Yes	HiSeq SR100

Lab ID	Sample type	Weight (mg)	Extraction	DSlib	SSlib	Capture	Sequencing type
MK46	Bone	41.4	Yes	Yes	Yes	Yes	MiSeq PE150/ HiSeq SR100
MK47	Bone	53.0	Yes	Yes	No	Yes	HiSeq SR100
MK48	Bone	61.4	Yes	Yes	Yes	Yes	MiSeq PE150/ HiSeq SR100
MK49	Bone	65.5	Yes	Yes	Yes	Yes	MiSeq PE75
MK50	Bone	68.0	Yes	Yes	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK51	Bone	47.1	Yes	Yes	Yes	Yes	HiSeq SR100
MK52	Bone	40.0	Yes	Yes	No	No	NA
MK53	Bone	45.0	Yes	No	Yes	Yes	HiSeq SR100
MK54	Bone	66.0	Yes	No	Yes	Yes	MiSeq PE75
MK55	Bone	45.0	Yes	No	Yes	Yes	MiSeq PE150/ HiSeq SR100
MK56	Bone	50.5	Yes	No	Yes	Yes	NA
MK57	Bone	30.0	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK58	Bone	48.7	Yes	No	Yes	Yes	HiSeq SR100
MK59	Bone	50.5	Yes	No	No	No	NA
MK60	Bone	35.1	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK61	Bone	17.2	Yes	No	No	No	NA
MK62	Bone	33.2	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE 75
MK63	Bone	20.5	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK64	Bone	58.3	Yes	No	Yes	Yes	HiSeq SR100
MK65	Bone	45.9	Yes	No	Yes	Yes	MiSeq PE75
MK66	Bone	43.5	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK67	Bone	52.0	Yes	No	Yes	Yes	HiSeq SR100
MK68	Bone	43.8	Yes	No	No	No	NA
MK69	Bone	28.1	Yes	No	Yes	Yes	HiSeq SR100
MK70	Bone	46.0	Yes	No	No	No	NA
MK71	Bone	37.1	Yes	No	Yes	Yes	HiSeq SR100
MK72	Bone	30.0	Yes	No	Yes	Yes	MiSeq PE150/ HiSeq SR100
MK73	Bone	49.0	Yes	No	Yes	Yes	HiSeq SR100
MK74	Bone	48.4	Yes	No	Yes	Yes	HiSeq SR100
MK75	Bone	44.0	Yes	No	Yes	Yes	HiSeq SR100
MK76	Bone	42.8	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK77	Bone	54.1	Yes	No	Yes	Yes	MiSeq PE75
MK78	Bone	56.2	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK79	Bone	45.8	Yes	No	No	No	NA
MK80	Bone	51.5	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75

Lab ID	Sample type	Weight (mg)	Extraction	DSlib	SSlib	Capture	Sequencing type
MK81	Bone	55	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK82	Bone	55	Yes	No	Yes	Yes	MiSeq PE75
MK83	Bone	49.7	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK84	Bone	54	Yes	No	Yes	Yes	MiSeq PE75
MK85	Bone	57	Yes	No	No	No	NA
MK86	Bone	47.3	Yes	No	Yes	Yes	MiSeq PE75
MK87	Bone	50	Yes	No	No	No	NA
MK88	Bone	50.3	Yes	No	No	No	NA
MK89	Bone	56.4	Yes	No	No	No	NA
MK90	Bone	43.7	Yes	No	No	No	NA
MK91	Bone	41.1	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK92	Bone	55.3	Yes	No	No	No	NA
MK93	Bone	51	Yes	No	No	No	NA
MK94	Bone	48.9	Yes	No	No	No	NA
MK95	Bone	41.3	Yes	No	No	No	NA
MK96	Bone	53	Yes	No	No	No	NA
MK97	Bone	53	Yes	No	No	No	NA
MK98	Bone	46.7	Yes	No	No	No	NA
MK99	Bone	48.5	Yes	No	No	No	NA
MK100	Bone	44.3	Yes	No	No	No	NA
MK101	Bone	51.8	Yes	No	No	No	NA
MK102	Bone	58.0	Yes	No	No	No	NA
MK103	Bone	46.4	Yes	No	Yes	Yes	MiSeq PE75
MK104	Bone	49.7	Yes	No	Yes	No	HiSeq SR100
MK105	Bone	57.1	Yes	No	No	No	NA
MK106	Bone	45.0	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK107	Bone	48.0	Yes	No	No	No	NA
MK108	Bone	41.9	Yes	No	Yes	Yes	MiSeq PE75
MK109	Bone	42.8	Yes	No	No	No	NA
MK110	Bone	53.7	Yes	No	Yes	Yes	HiSeq SR100
MK111	Bone	55.9	Yes	No	No	No	NA
MK112	Bone	56.1	Yes	No	Yes	Yes	HiSeq SR100
MK113	Bone	49.1	Yes	No	No	No	NA
MK114	Bone	49.1	Yes	No	No	No	NA
MK115	Bone	45.7	Yes	No	Yes	Yes	HiSeq SR100
MK116	Bone	37	Yes	No	Yes	Yes	HiSeq SR100
MK117	Bone	26.7	Yes	No	No	No	NA
MK118	Bone	57.5	Yes	No	Yes	Yes	HiSeq SR100
MK119	Bone	37.1	Yes	No	No	No	NA

Lab ID	Sample type	Weight (mg)	Extraction	DSlib	SSlib	Capture	Sequencing type
MK120	Bone	33.0	Yes	No	Yes	Yes	NA
MK121	Bone	16.9	Yes	No	Yes	Yes	HiSeq SR100
MK122	Bone	21.1	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK123	Bone	18.8	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK124	Bone	31.2	Yes	No	No	No	NA
MK125	Bone	45.3	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK126	Bone	53.1	Yes	No	No	No	NA
MK127	Bone	51.1	Yes	No	No	No	NA
MK128	Bone	52.4	Yes	No	Yes	Yes	MiSeq PE75
MK129	Bone	51.6	Yes	No	No	No	NA
MK130	Bone	48.7	Yes	No	Yes	Yes	MiSeq PE75
MK131	Oesophagus tissue	-	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK132	Oesophagus tissue	-	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK133	Body tissue	-	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK134	Toepad tissue	-	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK135	Toepad tissue	-	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK136	Feather	-	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK137	Feather	-	Yes	No	Yes	Yes	Nano MiSeq PE70/ MiSeq PE75
MK138	Toepad tissue	-	Yes	No	Yes	No	HiSeq SR80
LastGA2_ Heart	Heart tissue	-	Yes	Yes	No	Yes	HiSeq Various

Table S3.2 Lab process table for all samples collected as part of this study. Table includes information on sample type, weight used in extraction, which library build method was used, if hybridisation capture was used, and which type of sequencing was performed. Those highlighted indicate samples which ultimately passed the filtering settings and were used in the final analysis.

Sample ID	Median Age if known	Median Age	TempNet Category kya=thousand years ago
MK10	2850	0.00285	1kya-12kya
MK15	1400	0.0014	1kya-12kya
MK21	4000	0.004	1kya-12kya
MK25	4000	0.004	1kya-12kya
MK28	2000	0.002	1kya-12kya
MK38	13250	0.01325	>12kya
MK40	13950	0.01395	>12kya
MK46	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK47	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK48	Unknown but unlikely to be over 5000	Unknown	1kya-12kya
MK49	Unknown but unlikely to be over 5000	Unknown	1kya-12kya
MK50	202	0.000202	<250
MK51	196	0.000196	<250
MK72	1800	0.0018	1kya-12kya
MK73	1800	0.0018	1kya-12kya
MK74	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK75	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK76	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK78	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK82	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK83	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK84	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK86	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK91	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK103	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK104	5000	0.005	1kya-12kya
MK106	5000	0.005	1kya-12kya
MK108	4250	0.00425	1kya-12kya
MK110	4250	0.00425	1kya-12kya
MK115	3350	0.00335	1kya-12kya
MK116	4950	0.00495	1kya-12kya
MK121	4350	0.00435	1kya-12kya
MK122	2250	0.00225	1kya-12kya
MK125	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK130	Unknown but unlikely to be over 1000 years old	Unknown	250-500
MK133	200	0.0002	<250
MK134	200	0.0002	<250
MK135	173	0.000173	<250
MK136	200	0.0002	<250
MK138	200	0.0002	<250
LastGA2_Heart	173	0.000173	<250

Table S3.3 Age information for samples which were used in the BEAST and TempNet analysis.

						Results when mapped to GA_mtgenome_OSLO_KU158188						
Lab ID	**Library type SE, PE, or * (for both)**	**Total number of SE reads**	**Total number of pairs**	**Total number of retained reads**	**Average number of NTs in retained reads**	**Total number of hits (prior to PCR duplicate filtering)**	**Total number of hits vs. total number of reads retained**	**Fraction of hits that were PCR duplicates**	**Total number of hits (excluding any PCR duplicates)**	**Total number of unique hits vs. total number of reads retained**	**Estimated coverage from unique hits**	**Average number of aligned bases per unique hit**
MK10	PE		1145987	1168457	58.67	938703	0.80337	0.99456	5102	0.00437	17.602	57.91
MK13	PE		576586	595791	55.80	162526	0.27279	0.99386	998	0.00168	3.322	55.88
MK15	PE		467164	474690	52.41	273647	0.57648	0.98952	2867	0.00604	9.500	55.62
MK21	PE		1078089	1134756	50.18	57860	0.05099	0.91772	4761	0.00420	15.252	53.77
MK25	*	8454960	1645926	6009299	50.77	748444	0.12455	0.99249	5620	0.00094	15.342	45.82
MK26	SE	7005801		5199184	40.37	455	0.00009	0.11429	403	0.00008	1.406	58.54
MK28	*	15372548	6342089	19468085	51.99	310788	0.01596	0.90133	30664	0.00158	122.777	67.20
MK29	SE	10022700		7991112	42.03	642	0.00008	0.28505	459	0.00006	1.600	58.52
MK30	PE		808415	846355	54.48	1749	0.00207	0.33047	1171	0.00138	3.580	51.31
MK31	PE		158063	155088	47.84	197	0.00127	0.41624	115	0.00074	0.367	53.60
MK34	PE		426833	431557	54.16	16851	0.03905	0.98505	252	0.00058	0.863	57.48
MK37	SE	1181690		971375	42.59	1788	0.00184	0.94855	92	0.00009	0.279	50.87
MK38	SE	14789745		12607841	51.90	40653	0.00322	0.96298	1505	0.00012	6.388	71.24
MK40	*	3692259	1873507	4624647	51.24	101361	0.02192	0.91065	9057	0.00196	25.742	47.70
MK42	PE		628699	654160	57.30	33395	0.05105	0.98380	541	0.00083	2.074	64.33
MK44	PE		619588	649236	59.47	2943	0.00453	0.89976	295	0.00045	1.048	59.64
MK45	SE	9160386		7557907	46.05	106	0.00001	0.48113	55	0.00001	0.181	55.15
MK46	*	6697251	3601212	9292709	49.71	44531	0.00479	0.68923	13839	0.00149	43.774	53.09
MK47	SE	2309668		2210675	44.04	43726	0.01978	0.81414	8127	0.00368	23.401	48.33
MK48	*	5378635	2392050	6119108	43.55	575343	0.09402	0.79342	118854	0.01942	369.559	52.19
MK49	*	607157	1328601	1938752	52.93	1149304	0.59281	0.97466	29129	0.01502	95.219	54.86
MK50	*	1007261	492096	1502692	59.38	848108	0.56439	0.96013	33818	0.02250	175.194	86.95
MK51	SE	6442312		5225053	40.38	52353	0.01002	0.63908	18895	0.00362	64.831	57.59
MK53	SE	954018		595157	37.64	1893	0.00318	0.88378	220	0.00037	0.675	51.50
MK54	PE		759852	891455	58.89	29590	0.03319	0.99649	104	0.00012	0.345	55.69
MK55	*	1351278	1779892	1360629	85.82	4215	0.00310	0.93286	283	0.00021	0.727	43.10
MK57	PE		616943	576859	47.59	106	0.00018	0.00943	105	0.00018	0.311	49.68

						Results when mapped to GA_mtgenome_OSLO_KU158188						
Lab ID	**Library type SE, PE, or * (for both)**	**Total number of SE reads**	**Total number of pairs**	**Total number of retained reads**	**Average number of NTs in retained reads**	**Total number of hits (prior to PCR duplicate filtering)**	**Total number of hits vs. total number of reads retained**	**Fraction of hits that were PCR duplicates**	**Total number of hits (excluding any PCR duplicates)**	**Total number of unique hits vs. total number of reads retained**	**Estimated coverage from unique hits**	**Average number of aligned bases per unique hit**
MK58	SE	9838472		7430415	42.14	2483	0.00033	0.53000	1167	0.00016	4.030	57.96
MK60	PE		208549	194105	50.20	2415	0.01244	0.91718	200	0.00103	0.569	47.75
MK62	PE		111336	109888	53.19	25744	0.23427	0.99674	84	0.00076	0.243	48.48
MK63	PE		726857	700891	53.27	3181	0.00454	0.93052	221	0.00032	0.666	50.55
MK64	SE	6458031		4838037	41.01	1563	0.00032	0.38196	966	0.00020	3.084	53.59
MK65	PE		873614	919981	60.35	3065	0.00333	0.88907	340	0.00037	1.107	54.64
MK66	PE		309023	311687	51.69	8022	0.02574	0.94702	425	0.00136	1.159	45.78
MK67	SE	7260998		5226387	47.65	5020	0.00096	0.96733	164	0.00003	0.547	56.01
MK69	SE	3421051		2662747	46.87	1114	0.00042	0.95332	52	0.00002	0.176	56.81
MK71	SE	17614027		11738258	45.26	96099	0.00819	0.99487	493	0.00004	1.647	56.08
MK72	*	1900651	2033404	3134784	54.82	151240	0.04825	0.69118	46706	0.01490	147.925	53.16
MK73	SE	32330272		23293864	42.47	2450232	0.10519	0.98965	25359	0.00109	89.017	58.92
MK74	SE	6063592		4831238	40.63	22587	0.00468	0.86399	3072	0.00064	9.761	53.33
MK75	SE	4307305		3267171	39.57	43950	0.01345	0.94146	2573	0.00079	7.655	49.93
MK76	PE		384370	357506	47.25	23902	0.06686	0.84905	3608	0.01009	11.178	52.00
MK77	PE		1369136	1369627	57.12	125097	0.09134	0.99615	482	0.00035	1.593	55.48
MK78	PE		153456	144577	47.81	49095	0.33958	0.93955	2968	0.02053	8.707	49.24
MK80	PE		191309	188180	48.62	18736	0.09956	0.99408	111	0.00059	0.352	53.15
MK81	PE		524993	516514	59.09	2812	0.00544	0.81508	520	0.00101	1.697	54.76
MK82	PE		4538224	4471446	52.69	263178	0.05886	0.99237	2007	0.00045	6.772	56.63
MK83	PE		489039	502957	65.68	21346	0.04244	0.88251	2508	0.00499	8.540	57.15
MK84	PE		2756073	2795101	57.39	965141	0.34530	0.98962	10016	0.00358	33.601	56.31
MK86	PE		3925441	3935947	57.08	1201748	0.30533	0.99741	3109	0.00079	10.087	54.45
MK91	PE		178038	174894	51.74	90017	0.51469	0.95500	4051	0.02316	13.018	53.94
MK103	PE		599411	603360	52.67	9470	0.01570	0.40528	5632	0.00933	18.388	54.80
MK104	SE	9017013		8115882	51.60	2784572	0.34310	0.98867	31544	0.00389	146.081	77.73

					Results when mapped to GA_mtgenome_OSLO_KU158188							
Lab ID	**Library type SE, PE, or * (for both)**	**Total number of SE reads**	**Total number of pairs**	**Total number of retained reads**	**Average number of NTs in retained reads**	**Total number of hits (prior to PCR duplicate filtering)**	**Total number of hits vs. total number of reads retained**	**Fraction of hits that were PCR duplicates**	**Total number of hits (excluding any PCR duplicates)**	**Total number of unique hits vs. total number of reads retained**	**Estimated coverage from unique hits**	**Average number of aligned bases per unique hit**
MK106	PE		468952	472856	55.41	323508	0.68416	0.97194	9078	0.01920	30.350	56.11
MK108	PE		869643	876993	53.81	485179	0.55323	0.91049	43426	0.04952	136.617	52.80
MK110	SE	12818613		11671609	54.38	7157685	0.61326	0.99686	22452	0.00192	84.489	63.16
MK112	SE	5433398		3444422	42.94	2322	0.00067	0.51852	1118	0.00032	3.434	51.55
MK115	SE	5313081		4491377	44.22	108963	0.02426	0.79649	22175	0.00494	90.606	68.58
MK116	SE	6572626		5226835	39.70	369429	0.07068	0.91817	30232	0.00578	122.122	67.80
MK118	SE	5822506		3588232	40.84	83	0.00002	0.43373	47	0.00001	0.142	50.60
MK121	SE	8655497		7383179	48.15	298656	0.04045	0.95245	14201	0.00192	51.475	60.84
MK122	PE		502683	497894	49.60	343386	0.68968	0.96493	12042	0.02419	33.992	47.38
MK123	PE		624973	660170	58.19	1217	0.00184	0.85456	177	0.00027	0.599	56.81
MK125	PE		523729	532859	53.79	94063	0.17653	0.97315	2526	0.00474	8.219	54.61
MK128	PE		610041	594553	46.92	34	0.00006	0.02941	33	0.00006	0.118	59.97
MK130	PE		320431	305710	45.97	5483	0.01794	0.21029	4330	0.01416	12.517	48.52
MK131	PE		300754	277897	42.24	44118	0.15876	0.31327	30297	0.10902	74.396	41.21
MK132	PE		550631	555358	52.91	257053	0.46286	0.99080	2365	0.00426	6.222	44.15
MK133	PE		429392	434144	47.07	398068	0.91690	0.97802	8750	0.02015	23.041	44.20
MK134	PE		343766	333106	54.24	147070	0.44151	0.41303	86326	0.25915	288.624	56.12
MK135	PE		579992	581760	54.73	96607	0.16606	0.71257	27768	0.04773	88.901	53.73
MK136	PE		563635	554179	53.79	29361	0.05298	0.16893	24401	0.04403	67.829	46.66
MK137	PE		193090	220496	59.91	3757	0.01704	0.98190	68	0.00031	0.224	55.28
MK138	SE	10796460		9368986	46.60	113937	0.01216	0.97543	2799	0.00030	9.756	58.50
LastGA2_Heart	SE	957970612		951654295	55.64	294476	0.00031	0.58609	121886	0.00013	430.094	59.22

Table S3.4 Summary data for mitogenome samples from PALEOMIX. Library type: PE= Paired-end, SE= Single end, *= both. Data used for molecular preservation comparison.

Lab ID	Genbank accession number
MK10	MF776845
MK15	MF776853
MK21	MF776854
MK25	MF776855
MK28	MF776856
MK38	MF776857
MK40	MF776858
MK46	MF776859
MK47	MF776860
MK48	MF776861
MK49	MF776862
MK50	MF776863
MK51	MF776864
MK72	MF776865
MK73	MF776866
MK74	MF776867
MK75	MF776868
MK76	MF776869
MK78	MF776870
MK82	MF776871
MK83	MF776872
MK84	MF776873
MK86	MF776874
MK91	MF776875
MK103	MF776841
MK104	MF776842
MK106	MF776843
MK108	MF776844
MK110	MF776846
MK115	MF776847
MK116	MF776848
MK121	MF776849
MK122	MF776850
MK125	MF776851
MK130	MF776852
MK133	MF188884
MK134	MF188885
MK135	MF188886
MK136	MF188887
MK138	MF188888
LASTGA2	MF188889

Table S3.5 Genbank accession numbers for samples used in analysis.

S3 Figures

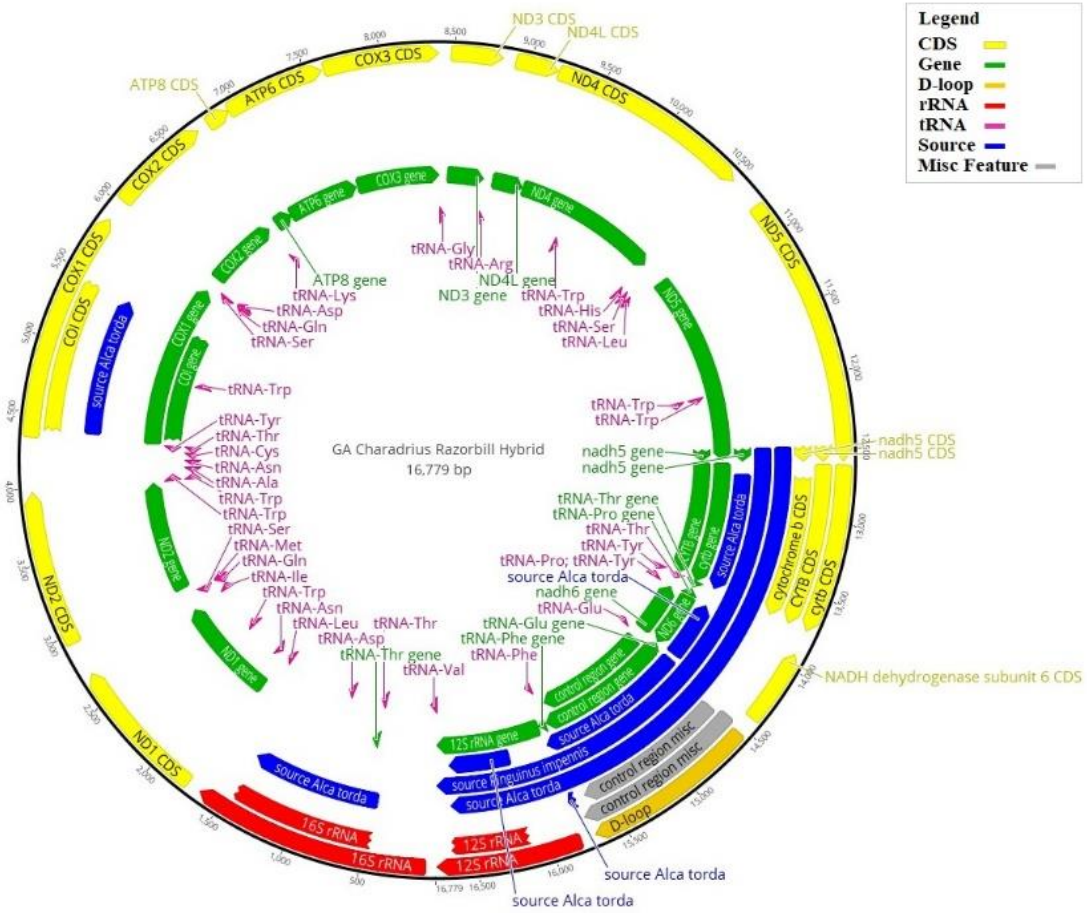

Figure S3.1 Illustration of the hybrid reference mitogenome constructed using the killdeer (*Charadrius vociferous*) mitogenome, with orthologous gene regions replaced by those of the great auk (*Pinguinus impennis*), or razorbill (*Alca torda*), when great auk data was unavailable. Annotations correspond to the various regions of the mitogenome, those in blue show where great auk or razorbill genes have been used, yellow corresponds to coding regions, green shows all gene regions, the D-loop is shown in gold, rRNA regions are in red, tRNA regions are in pink and any miscellaneous features are in grey. The numbers on the outer black circle correspond to the base position of the mitogenome.

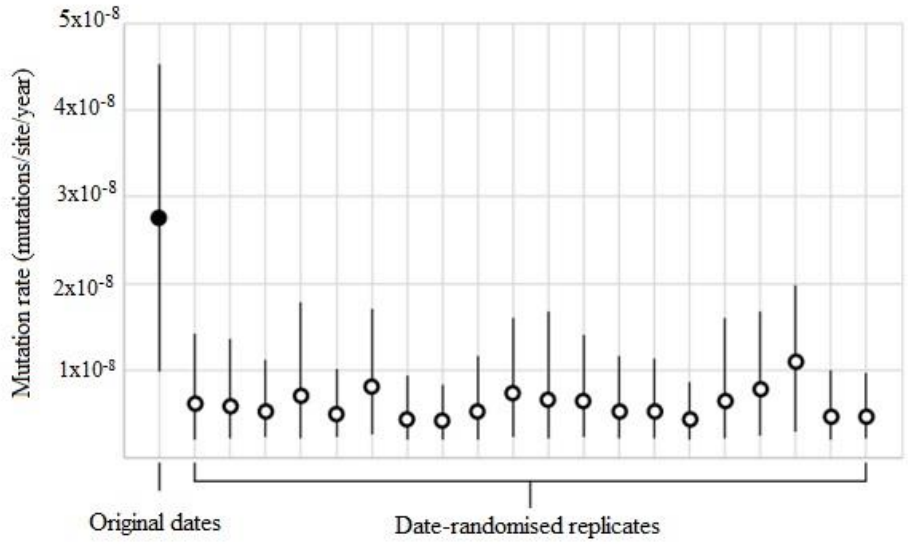

Figure S3.2 Date-randomization test for temporal structure in mitogenome sequences from 25 individuals. The filled circle indicates the median posterior estimate of the mutation rate from the original data set, whereas the empty circles show the median posterior estimates from 20 date-randomised replicate data sets. The 95% credibility intervals (vertical black lines) of the estimates from the date-randomised replicates do not overlap with the median estimate from the original data set, providing some evidence of temporal structure in the data set (criterion CR1). However, the 95% credibility intervals of the estimates from the date-randomised replicates overlap with the 95% credibility interval of the estimate from the original data set, indicating that the data set does not meet the more stringent criterion CR2.

S3 Additional information

S3.1 Method comparison

S3.1.1 Introduction

Within the field of aDNA methods are continually being developed to improve the success of obtaining and sequencing the smallest and most degraded fragments of DNA, and to increase the endogenous DNA content. To determine the most appropriate and effective methods for the samples in this study, a preliminary experiment was performed in which we compared library build methods, as well as the effectiveness of using hybridisation enrichment capture for whole mitochondrial genomes versus shotgun sequencing.

Each method is said to have various advantages and disadvantages. For example, the Single-Stranded library (SSlib) preparation method developed by Gansuage & Meyer (2013) was developed specifically for aDNA samples and therefore takes into account the characteristics of aDNA samples, such as the highly fragmented nature of the DNA (SSlib preparation consistently recovers very short DNA molecules <50bp) and that target DNA is present in low amounts. The SSlib build method is reported to increase library yields, library complexity and endogenous DNA content (Gansauge & Meyer, 2013; Wales *et al.*, 2015). However, the SSlib build method is more labour intensive (~4h for DSlib, ~9h for SSlib as reported in Wales *et al.*, 2015) and expensive, than the DSlib build method by Meyer & Kircher (2010) (Wales *et al.*, 2015).

In addition to comparing library build methods, the use of hybridisation enrichment capture was also assessed. This method is used to enrich libraries for target DNA, for example in the case of this study, for great auk mitogenomes. It works by 'capturing' the target DNA from the library using target specific bait and immobilising this onto magnetic beads, this then allows for the non-target/contamination DNA (characteristic of aDNA samples) to be washed away. This method is said to increase the efficiency of sequencing, however, there are disadvantages and limitations of this method too (Knapp & Hofreiter, 2010).

S3.1.2 Methods

Following DNA extraction using the Dabney *et al.* (2013) method (See 3.2.1), eight samples were selected for double strand library (DSlib) build (Meyer & Kircher, 2010) (Table S3.6). These samples were sent for shotgun sequencing (Illumina). In light of the shotgun sequencing results, a second library build method was tested in addition to using hybridisation enrichment

capture. Three of the original samples used in the DSlib shotgun experiment were then used in hybridisation capture to enrich for great auk mitochondrial DNA using MYcroarray MYbaits Sequence Enrichment protocol v2.3.1 (MYcroarray MYbaits, 2014). Another set of samples were used for the single-strand library build method (Gansauge & Meyer, 2013 with modifications described in Bennet *et al.*,2014). As before, the samples were sent for sequencing both with and without hybridisation capture (Table S3.6).

Sample	Country	DSlib Shotgun or Capture	SSlib Shotgun or Capture
MK10	Cladh Hallan, South Uist, Scotland	DSlib Shotgun	
MK17	Bornais Mound 1, South Uist, Scotland	DSlib Shotgun	
MK18	Bornais Mound 1, South Uist, Scotland	DSlib Shotgun	
MK24	Tofts Ness, Sanday, Orkney, Scotland	DSlib Shotgun	
MK25	Tofts Ness, Sanday, Orkney, Scotland	DSlib Shotgun	SSlib Shotgun SSlib Capture
MK26	Royal Manor Field, Portland, England	DSlib Shotgun	
MK28	Royal Manor Field, Portland, England	DSlib Shotgun DSlib Capture	SSlib Shotgun SSlib Capture
MK40	Santa Catalina, Lekeitio, Spain	DSlib Shotgun	SSlib Shotgun SSlib Capture
MK46	Funk Island, Newfoundland, Canada	DSlib Capture	SSlib Shotgun SSlib Capture
MK48	Vardø, Norway	DSlib Capture	SSlib Shotgun SSlib Capture
MK55	Frederiksborg, Denmark		SSlib Shotgun SSlib Capture
MK72	Schipluiden, Netherlands		SSlib Shotgun SSlib Capture

Table S3.6 Sample information for samples used in method comparison experiments and details of library build method and use of capture for samples used in comparison experiment.

S3.1.3 Results

S3.1.3.1 Double-stranded library with shotgun sequencing

The results of the double-stranded library build with shotgun sequencing showed that this method produced very low results with respect to percentage of unique reads mapping. The percentage of unique reads mapping ranged from 0.0001% (MK28) to 0.0293% (MK10) (Table S3.7).

	Total number of raw reads				Total number of unique reads mapping			
Sample	**DSlib Shotgun**	**DSlib Capture**	**SSlib Shotgun**	**SSlib Capture**	**DSlib Shotgun**	**DSlib Capture**	**SSlib Shotgun**	**SSlib Capture**
MK10	23,911				7			
MK17	510,316				4			
MK18	234,520				4			
MK24	3,279,683				6			
MK25	148,369		1,371,696	274,230	4		122	3,710
MK26	1,316,221				4			
MK28	766,591	3,777,461	5,423,843	918,246	1	76	27	2,253
MK40	2,000,416		1,148,814	724,693	13		30	4,969
MK46		1,842,620	2,217,985	754,018		2,198	35	2,603
MK48		544,315	1,441,981	950,069		2,889	607	88,861
MK55			1,523,618	256,274			186	129
MK72			1,070,955	456,545			213	25,723

	Percentage of unique reads mapping			
Sample	**DSlib Shotgun**	**DSlib Capture**	**SSlib Shotgun**	**SSlib Capture**
MK10	0.0293			
MK17	0.0008			
MK18	0.0017			
MK24	0.0002			
MK25	0.0027		0.009	1.353
MK26	0.0003			
MK28	0.0001	0.002	0.001	0.245
MK40	0.0006		0.003	0.686
MK46		0.119	0.002	0.345
MK48		0.531	0.042	9.353
MK55			0.012	0.050
MK72			0.020	5.634

Table S3.7 Result for comparisons between library build method and shotgun sequencing vs. capture.

S3.1.3.2 Double-stranded library: shotgun sequencing vs. capture

Sample MK28 was the only sample for which comparison between shotgun sequencing and sequencing following hybridisation capture with double-stranded library build could be performed. The result of the comparison showed that the percentage of unique reads mapping increased from 0.0001% (shotgun) to 0.002% following capture (Table S3.7), a 15times increase. However, despite not being used in the DSlib and shotgun experiment, we saw that

when samples MK46 and MK48 were used in DSlib capture, the percentage of reads mapping was 0.119% and 0.531%, respectively, which when compared to the overall results of non-captured DSlib samples as discussed above, shows the overall improvement to percentage of reads mapping that hybridisation capture has made.

S3.1.3.3 Double-stranded library with shotgun sequencing vs. single-stranded with shotgun sequencing

Comparisons were made for three samples to assess the effect of library build method when samples were shotgun sequenced. For sample MK25 0.0027% of unique reads mapped with the double-stranded library build method, whereas, 0.0089% mapped following the single-stranded library build method. With sample MK28, we again saw an increase in the percentage of reads mapping with the single-stranded library build method, with 0.0001% of reads mapping with DSlib build and 0.0005% mapping following SSlib build. This is again shown with sample MK40 for which 0.0006% unique reads mapped with the DSlib build and 0.0026% with the SSlib build method (Table S3.7).

S3.1.3.4 Double -stranded library with capture vs. single-stranded with capture

Three samples were compared for library build method following capture. Samples MK28, MK46 and MK48 all showed an increase in the percentage of unique reads mapping with the single-stranded library prep method (MK 28= 0.002% (DSlib) to 0.25% (SSlib), MK46=0.1193% (DSlib) to 0.35% (SSlib), MK48= 0.53% (DSlib) to 9.35% (SSlib) (Table S3.7).

S3.1.3.5 Single-stranded library: shotgun sequencing vs. capture

For the comparisons between samples for which the single-stranded library build method had been used, we were able to compare 7 samples for shotgun vs capture. For the single-stranded library samples with shotgun sequencing, the percentage of unique reads mapping ranged from 0.0005% (MK28) to 0.0421% (MK48). When hybridisation capture was used, the results for percentage of unique reads mapping ranged from 0.05% (MK55) to 9.35% (MK48). Performing hybridisation enrichment capture caused an average increase of percentage of unique reads mapping of 2.51% (range 0.04% (MK55) to 9.31% (MK48)) (Table S3.7).

S3.1.3.6 Double-stranded library with capture vs. Single-stranded with shotgun sequencing

Comparisons between DSlib prep with capture and SSlib prep with shotgun, highlighted the benefit of using hybridisation capture. The results of DSlib with capture for MK28, MK46 and MK48 were 0.002%, 0.119% and 0.531%, respectively, whereas with SSlib and shotgun they were lower at 0.001%, 0.002% and 0.042% (Table S3.7). This comparison therefore shows that despite the rest of the comparisons highlighting that the SSlib preparation method was more effective at increasing the number of reads mapping, when capture was used with the DSlib prep this gave a better result than SSlib without capture, showing the value of using hybridisation capture.

S3.1.3.7 Double-stranded library with shotgun sequencing vs. Single-stranded library with capture

The final comparison made to confirm the effectiveness of the single-stranded library build method combined with hybridisation capture vs. the original double-stranded library build method with shotgun sequencing was performed for three samples. For MK25 the percentage of unique reads mapping with DSlib build and shotgun was 0.0027%, and with SSlib build with capture, 1.35%. Sample MK28 had 0.0001% reads mapping following DSlib build and shotgun but 0.25% mapping following SSlib build and capture. Finally, MK40 had 0.0006% reads mapping for DSlib shotgun and 0.69% with SSlib and capture (Table S3.7). This showed that the use of SSlib build with capture can increase the number of unique reads mapping by at least 502 times (MK25) but up to 1881 times (MK28).

S3.1.4 Conclusion

The overall result of the comparisons between the various library build methods and the effectiveness of using hybridisation enrichment capture was that using the single-stranded library build method, combined with hybridisation enrichment gave the best results, in terms of percentage of unique reads mapping. This result is in agreement with previous published research. Therefore, a SSlib build method, combined with hybridisation enrichment, was chosen to be used for the remaining samples of the study in hope that this would provide us with the best results.

S3.2 Demographic inferences using 'strict' alignment

We reconstructed changes in population size over time using the BEAST v1.7.5 and v1.8.4 (Drummond *et al.*, 2012) software packages. Details of substitutions rates and tip calibration can be found in the main text. As with the 'relaxed' alignment, for the strict alignment, we compared strict vs. uncorrelated relaxed lognormal clock and, independently, a constant size coalescent tree prior vs. the more flexible Bayesian skyline plot (BSP) (with 10 groups). In all cases, in order to reduce error introduced by more parameter rich models, these more complex models (relaxed clock; BSP) were only accepted if they were 'decisively' better than the respective simpler models (Kass & Raftery, 1995). For each BEAST analysis, we ran three independent Markov Chain Monte Carlo (MCMC) chains for 20 million generations, sampling trees and model parameters every 2,000 generations, with the first 10% of each chain discarded as burn-in. We compared the results from each run in Tracer v1.5.0 to confirm convergence of the MCMC chains and achieving at least 200 as effective sample size (ESS) for each parameter (Drummond and Bouckaert, 2015).

Bayes factor (BF) model comparisons showed that the 'strict' alignment was best described by the simpler strict molecular clock mode, as it was not 'decisively' worse than the more complex relaxed lognormal clock model. When the BSP was compared with the constant population size model, the BF support showed the BSP to be 'decisively' better than the simpler constant population size model, showing an increase in effective population size over time (Fig. S3.3).

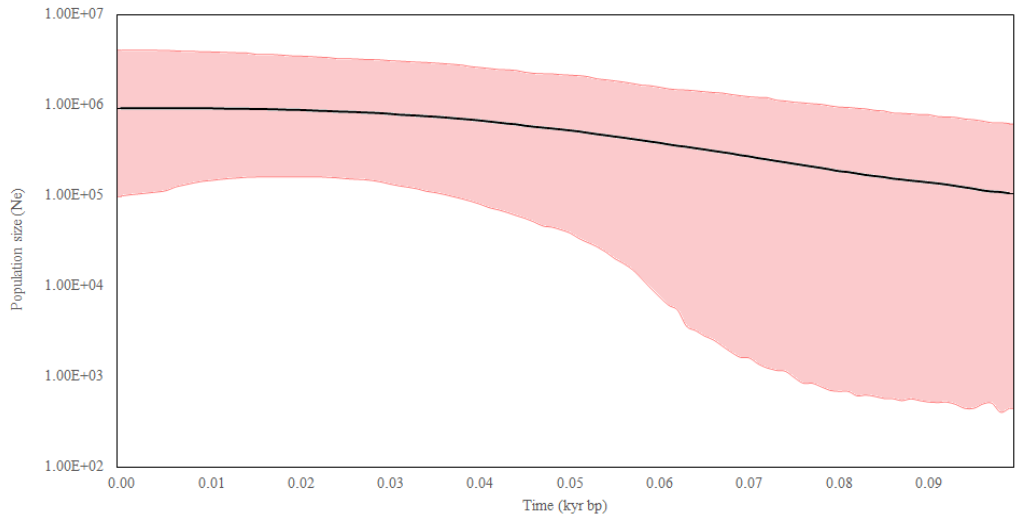

Figure S3.3 Bayesian skyline plot of temporal changes in female effective population size, created using the strict alignment, with a strict molecular clock. The solid black line indicates the mean value and the shaded area shows the 95% highest posterior density interval.

S3.3 Alternative TempNet grouping

Genetic diversity through space and time was visualised using statistical parsimony and a temporal haplotype network as implemented in TempNet (Prost & Anderson, 2011). To test the influence of sample groupings on the TempNet analysis, an alternative group configuration, to the one presenting in the main text (Fig. 3.3), was tested. Samples were divided into the following 4 categories determined by the age distribution of our samples: less than 1000 years old, 1000-4000 years old, 4000-10,000 years, and greater than 10,000 years old (Fig. S3.4).

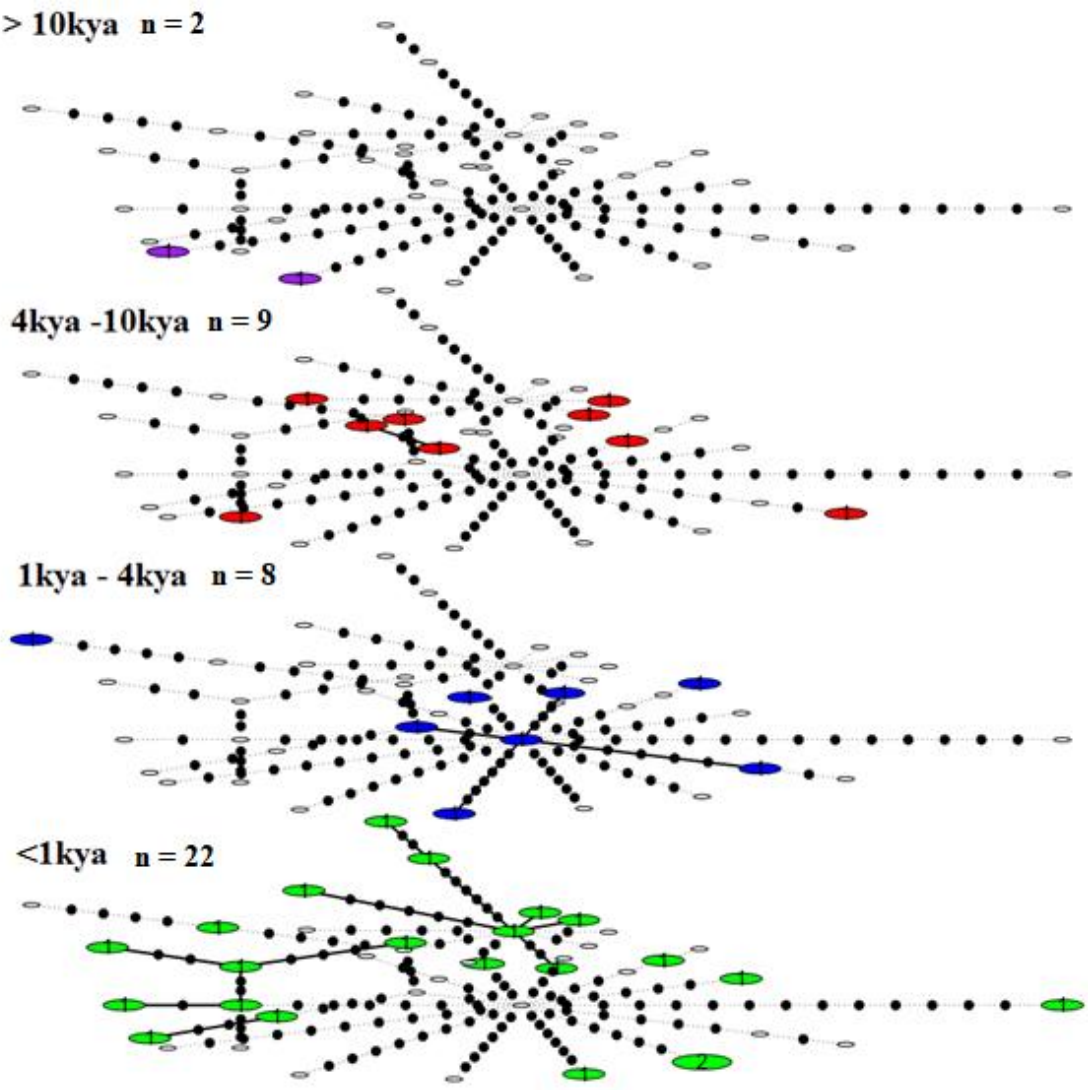

Figure S3.4 TempNet showing the haplotype diversity through time. Samples were split into four age categories. Each age category is represented by the same statistical parsimony network, with haplotypes present in this time category shown in colour, haplotypes present in another time category shown as empty dots, and mutation between haplotypes marked as filled dots. All samples have been included in this figure. For samples which did not have dates, BEAST was used to estimate their date and they were entered into the appropriate category.

S3.4 Population viability analysis- additional information and justification

Age of first breeding for the great auk is estimated to be between 4-7 years old (Bengtson, 1984), and a value of 5 years was therefore adopted for the model. Using this age makes analyses more conservative, as the younger the age of first breeding, the less susceptible to extinction the species is. The species is assumed to have been monogamous, with both sexes having had a single brood patch, suggesting only one egg was laid per breeding season, which occurred only once a year (Bengtson, 1984). Life expectancy is estimated to have been between 20-25 years (Bengtson, 1984) and we assume breeding until death. As several alcid species breed annually once they reach sexual maturity (De Santo and Nelson, 1995), we set reproductive rate to a conservative rate of 90% adult females breeding. With regards to mortality rates, this information was also estimated based on records from extant alcids. De Santo & Nelson (1995) report survival rates for alcid species at various life stages. Mortality at age 0-1 includes hatchlings and fledglings. For the great auk, we estimate mortality to be 30% due to the hatching success rate of puffins (*Fratercula arctica*) 72% and razorbill 78% (however, they may lay a replacement egg), and fledgling success of razorbill at over 90% and puffins at 73% (De Santo and Nelson, 1995). With regards to the model, juvenile mortality would cover mortality from age groups 1-2, 2-3, 3-4 and 4-5; therefore, our juvenile mortality rate was divided between these groups. Juvenile survival in razorbills is reported to be 32% , in common murre (*Uria aalge*) 30%, in thick-billed murre (*Uria lomvia*) 34% and in black guillemot (*Cepphus grylle*) 27% (De Santo and Nelson, 1995). Therefore, it is likely that the juvenile mortality rate is ~70%, which is the figure that we used for great auks, divided between the four age categories to give 1-2= 18%, 2-3= 18%, 3-4= 17%, 4-5= 17%. Annual adult survival rate is estimated to be quite high for great auks due to their large size (Bengtson, 1984; Montevecchi & Kirk, 1996). Annual adult survival in other alcids is also high, (razorbill 90% (De Santo and Nelson, 1995)), therefore we use an annual mortality rate of 10% for adult great auks.

S3.5 Molecular preservation

It is well documented that the preservation of DNA in ancient samples can vary greatly. Factors impacting molecular preservation include sample type/material and age. Here we have looked in more detail at the average number of aligned bases to the reference genome and the estimated

coverage of unique hits, to see if this shows any important or interesting information regarding molecular preservation of great auk mitochondrial DNA in the samples used in this study.

S3.5.1 Sample summary

In this study, the majority of samples used were bones, but samples from mounted museum specimens and preserved organs were also collected and processed. Samples ranged in age from 173 years old to ~15,000 years old. The mean average number of aligned bases to the reference mitogenome for all samples was 55.12bp, with a range of 41.21-86.95bp. Sample MK131, a 173-year-old tissue sample stored in spirit from Iceland, had the shortest average read length mapping at 41.21bp, and MK50, a bone sample removed from a mounted specimen killed in Iceland in 1821, had the longest average read length mapping at 86.95bp. The mean estimated coverage from unique hits, was 41.23x, with a range of 0.12-430.09x. The sample with the poorest estimated coverage was MK118 (0.12x), a bone sample of approximately 7000-9000 years old from Norway. The sample with the highest estimated coverage was LASTGA2_Heart (430.09), a tissue sample stored in spirit collected from the last pair killed on Eldey Island, June 1844. As sample LASTGA2_Heart has had a greater level of sequencing, we also looked at the results without this sample included. When LASTGA2_Heart was excluded, sample MK48, a bone from Norway of unknown age, had the highest estimated coverage (369.56x) and the overall mean estimated coverage was 35.98x. The mean average read length was 55.06bp.

S3.5.2 Comparisons based on sample type

The samples sent for sequencing can be split into four groups based on sample type: bones (n=66), tissue samples stored in spirit (n=3 or n=2), tissue from mounted specimen (n=4) and feather (n=2). Table S3.8 shows the results for the average read length mapping to the reference genome of unique reads, over 30bp, and the range of the average read length. It also shows the mean and range estimated coverage from unique hits. This information is also displayed visually in Fig. S3.5 and Fig. S3.6.

For bone samples, the average length of reads mapping to the reference mitogenome ranged from 43.10– 86.95bp, with an overall average of 55.68bp. For tissue samples taken from mounted museum specimens, the range was 44.20-58.20bp, with an average length mapping of 53.14bp. The feather samples had an average read length mapping to reference mitogenome

range of 44.66-55.28bp, with an average of 50.97bp. Tissue samples which had been stored in spirit had the lowest average read length mapping, 48.2bp and a range of 41.21-59.22bp. In terms of the estimated coverage from unique hits, for bone samples, the range was 0.12-369.56x, with an average of 31.87x. For tissue samples stored in spirits it ranged from 6.22-430.09x, when LASTGA2_Heart was included, and 6.22-74.40x without. The range for tissue from mounted specimens was 9.76-288.62x, with an average of 102.58x. Finally, for feathers it ranged from 0.22-67.83x.

Sample type	Number of samples	Mean average read length mapping to reference mitogenome (bp)	Average read length mapping to reference mitogenome (bp) range	Mean estimated coverage from unique hits	Estimated coverage from unique hits range
Bone	66	55.68	43.10-86.95	31.87	0.12-369.56
Tissue stored in spirit	3	48.2	41.21-59.22	170.24	6.22-430.09
Tissue stored in spirit (no heart)	2	42.68	41.21-44.15	40.31	6.22-74.40
Tissue from mounted specimen	4	53.14	44.20-58.50	102.58	9.76-288.62
Feather	2	50.97	44.66-55.28	34.03	0.22-67.83

Table S3.8 Data table for comparisons between sample type.

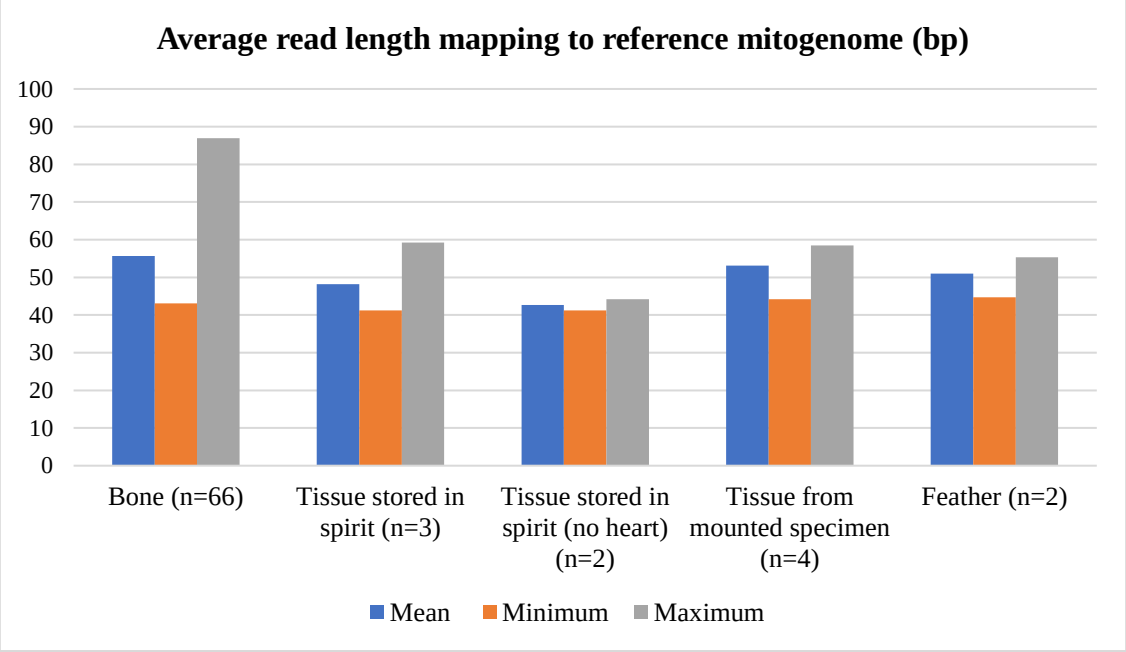

Figure S3.5 Comparisons of average read length mapping to the reference mitogenome for each of the sample type groups. The minimum, maximum and mean read length have been displayed.

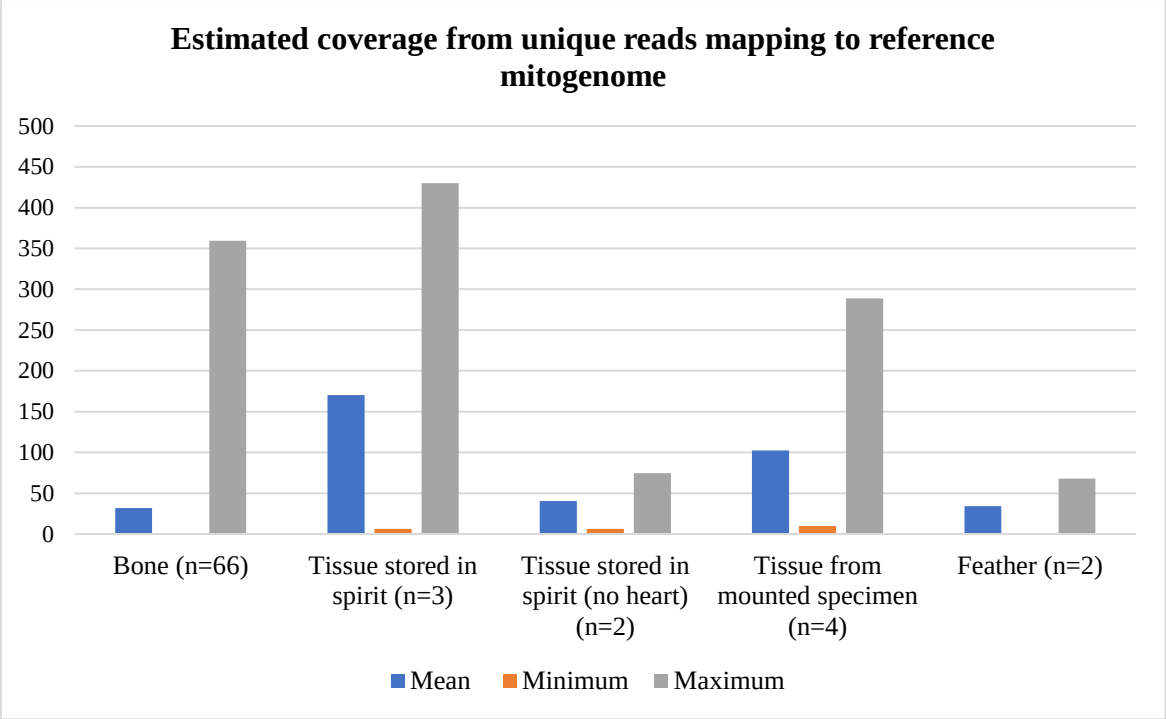

Figure S3.6 Comparisons of estimated coverage from unique reads mapping to the reference mitogenome, for each of the sample type groups. The minimum, maximum and mean estimated coverage is displayed.

S3.5.3 Comparisons based on sample age

Samples were also divided into age groups, based on the same justification used for the TempNet analysis (3.2.6). Samples less than 250 years had an average read length mapping to the reference genome of 54.87bp, with a range of 41.21-86.95bp. The estimated coverage of unique hits ranged from 0.22-430.09x (with LASTGA2_Heart) and 0.22-288.62x (without LASTGA2_Heart). Samples aged between 250-1000 years old had an average read length of 54.02bp, with a range of 48.33-59.97bp (Table S3.9, Fig. S3.7). Estimated coverage from unique hits ranged from 0.12-43.77x, with an average of 10.56x. Samples over 1000 but less than 12000 contained 38 of the samples. In this group the average read length was 55.32bp, with a range of 43.1-77.73bp. The estimated coverage was 0.14-369.56x, with an average of 42.27x. The last group included samples over 12,000years old. In this group the average read length ranged from 47.7-71.24, with an average of 58.76. The estimated coverage of unique hits ranged from 0.28-25.74x, with an average of 7.12x (Table S3.9, Fig. S3.8).

Age	Number of samples	Mean average read length mapping to reference mitogenome (bp)	Average read length mapping to reference mitogenome (bp) range	Mean estimated coverage from unique hits	Estimated coverage from unique hits range
<250	11	54.87	41.21-86.95	111.73	0.22-430.09
<250 without heart	10	54.44	41.21-86.95	79.9	0.22-288.62
>250-<1000	21	54.02	48.33-59.97	10.56	0.12-43.77
>1000-12000	38	55.32	43.1-77.73	42.27	0.14-369.56
>12000	5	58.76	47.7-71.24	7.12	0.28-25.74

Table S3.9 Data table for comparisons of between age groups.

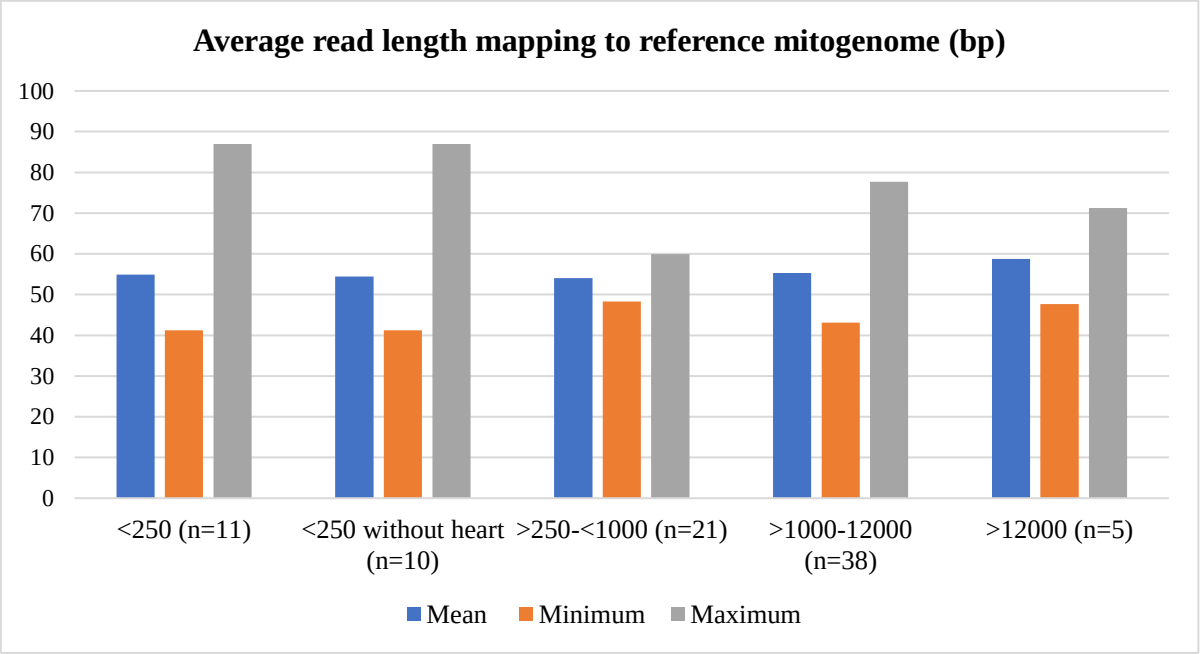

Figure S3.7 Comparisons of average read length mapping to the reference mitogenome for each of the age groups. The minimum, maximum and mean read length have been displayed.

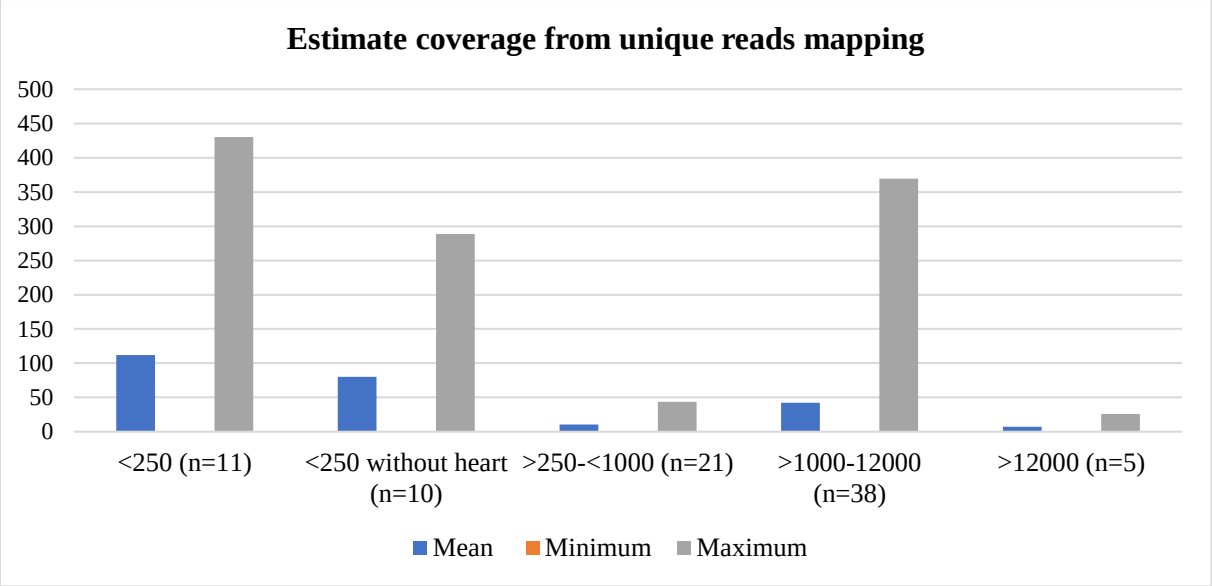

Figure S3.8 Comparisons of estimated coverage from unique reads mapping to the reference mitogenome, for each of the age groups. The minimum, maximum and mean estimated coverage is displayed.

S3.5.4 Discussion

As molecular preservation is an important and interesting component of aDNA studies, it was useful to perform the above comparisons, however, any conclusions that may be drawn from the data is limited for a number of reasons. Firstly, there is a skewed distribution of samples in the various comparative groups (e.g. n=66 bone, n=2 feather). Secondly, the comparisons made between average coverage are difficult to interpret as it relies on complete evenness of libraries in the sequencing run, and this is never the case. Thirdly, as samples were not shotgun sequenced, we were unable to compare levels of endogenous content, which would have been the most informative comparison for molecular preservation. Therefore, in the absence of this, the most unambiguous measure for sample preservation may be fragment length. However, this also has issues. Here we have reported the average fragment length of reads mapping to the reference genome. This means that some reads have already been discarded during read filtering (i.e. reads less than 30bp have been discarded and PCR duplicate reads removed). Nevertheless, the comparisons which have been reported above regarding fragment length do not show a clear trend between sample age nor type. Finally, DNA preservation is dependent on a large number of factors which differ widely for each individual sample. This makes it

difficult to draw any reliable conclusions about sample preservation and any one factor such as age or tissue type, as the other variables may have impacted preservation.

S3.6 GPS-equipped drifting capsules

Following release, easterly winds prevailed and the two GPS-equipped drifting capsules drifted westwards past the tip of the Reykjanes peninsula and past Eldey Island (Fig. 3.5 main text). Over the next two weeks, the capsules drifted towards Greenland and when located near the continental shelf started drifting southwards along the coast. The capsules then followed the track of the Icelandic Low, a low-pressure area found between Iceland and Southern Greenland in winter. The Icelandic Low took the capsules in an anti-clockwise circle back towards Iceland, and onward again towards the west coast of Greenland. In summer, the Icelandic Low weakens so in late April the capsules turned westwards past the southern tip of Greenland and into the Labrador Sea. In summer, they drifted slowly towards the Labrador coast until the beginning of August when they started drifting south-eastwards along the coast of Labrador and Newfoundland and past Funk Island and around 500km east. By the end of October, the capsules start to follow the trail of the winter low pressures across the Atlantic. In the beginning of January, capsule 1 drifted eastwards, around 50km south of St. Kilda and came ashore on the island of Tiree (15.01.2017). Capsule 2 drifted northwards passing around 70km west of St. Kilda and west of the Faeroes towards Iceland. In early March, the capsule was around 20km from the east coast of Iceland when it turned eastwards and then towards south by the beginning of April. It drifted towards the Faeroes where it came ashore on the island of Sandoy (13.05.2017).

The forces driving the capsules are currents, wind and waves. The capsules got trapped in the Iceland Low where the wind direction is in a counter clockwise circle in winter in the Denmark strait. In spring, when the Iceland Low starts dissolving, they pass Cape Farewell and in summer, they drift slowly in calmer summer winds and followed the cold current towards the Labrador coast and then along the coast of Labrador and Newfoundland. In autumn, they hit the path of lows crossing the Atlantic as well as following the warmer Gulfstream. In spring, when at the east coast of Iceland, the weather was calmer and thus capsule 2 drifted slowly towards and then away from the coast and ending up in the Faeroes.

Chapter 4: Exploring the Possibility of Using Nuclear DNA Sequencing to Characterise the Population Genetics of the Great Auk

Tentative author list:

Jessica E. Thomas, Filipe J. G. Vieira, Jose A. Samaniego Castruita, Michael Hofreiter, John R. Stewart [†], Gary R. Carvalho[†], M. Thomas P. Gilbert[†] and Michael Knapp[†] (Equal contribution [†])

Abstract

For much of the history of ancient DNA research, studies were dominated by mitochondrial DNA (mtDNA) markers. However, advancements to laboratory protocols and sequencing technologies have greatly improved the success of recovering nuclear DNA (nuDNA). It is now generally accepted that using a combination of mtDNA and nuDNA markers can provide a more nuanced insight when attempting to answer evolutionary and ecological questions. The results of our mitochondrial genome study (Chapter 3) revealed a lack of population genetic structure and high genetic diversity in the great auk. It would consequently be useful to use nuDNA to further investigate these results, as this would provide us with a more detailed picture of great auk evolution and extinction. Initially, twenty-four samples were chosen for capture of 495 nuclear markers, twelve of which were screened by sequencing to determine whether the capture was successful. Unfortunately, coverage of the 495 targeted markers in the twelve samples was extremely low. Only one sample, MK134, had any genes with at least 3-fold coverage. Therefore, further analysis that would provide any meaningful results could not be performed. Consequently, the discussion of this chapter focuses on the possible reasons for why the results were poor, in light of offering suggestions on ways to improve the success of targeting nuDNA from great auk samples in the future.

4.1 Introduction

In the early days of ancient DNA (aDNA) and modern DNA research, the majority of population genetic studies focused on using Polymerase Chain Reaction (PCR) to amplify short sections of mitochondrial DNA (mtDNA), especially barcoding genes such as *12S rRNA, CO1, Cytb* and the control region (Zhang & Hewitt, 2003; Hagelberg *et al.*, 2014; Grealy *et al.*, 2017a). As a result of advancements in sequencing technology, i.e. the introduction of 'Next Generation Sequencing' (NGS), whole genomes have now been sequenced from hundreds of species (e.g. Genome 10K project-(Genome 10K Community of Scientists, 2009; Bernardi *et al.*, 2012; Jarvis *et al.*, 2014; Zhang, C. Li, *et al.*, 2014; Pan *et al.*, 2016)) and modern population genetic studies now routinely include whole genome data (e.g. (Francioli *et al.*, 2014). These technological improvements have also opened the door to palaeogenomics (Hagelberg *et al.*, 2014; Der Sarkissian *et al.*, 2015). Nevertheless, aDNA population studies are still dominated by mitochondrial genome data.

The initial dominance of mtDNA use in the field of aDNA was due to the difficultly of obtaining useable nuclear DNA (nuDNA) from ancient samples (Binladen *et al.*, 2006; Rohland *et al.*, 2009; Grealy *et al.*, 2017a). Mitochondrial DNA is typically easier to sequence successfully due to a number of reasons. For example, mtDNA is present in higher copy numbers than nuDNA (Huynen *et al.*, 2003) (typically 1000 times more mtDNA than nuDNA (Rizzi *et al.*, 2012)) and therefore more likely to be amplified (Der Sarkissian *et al.*, 2015; Grealy *et al.*, 2017a). Nuclear DNA has also been argued to degrade at least twice as fast as mtDNA in at least some situations (Allentoft *et al.*, 2012). However, advancements in laboratory protocols and the reduced cost of sequencing has made sequencing nuDNA from aDNA samples possible. This is evident from the increasing number of publications that present nuclear gene fragments (e.g. moa (Huynen *et al.*, 2003), sloth (Poinar *et al.*, 2003), mammoth (Greenwood *et al.*, 1999)) and even draft genomes (e.g. mammoth (Miller *et al.*, 2008), Palaeo Eskimo (Rasmussen *et al.*, 2010), Denisovans (Meyer *et al.*, 2012), passenger pigeon (Hung *et al.*, 2014; Murray *et al.*, 2017)) from a number of ancient specimens.

Although mtDNA data has some factors that make it suitable and desirable for studying population genetics (Shields & Helm-Bychowski, 1988), such as sensitivity for detecting population bottlenecks and restrictions in geneflow (Shields & Helm-Bychowski, 1988; Friesen *et al.*, 2007), these same traits deemed good for use as a molecular marker, can also be

a limitation. For example, one of the main features of mtDNA is that it is inherited maternally. This is good as mtDNA is essentially maternal clones, and therefore differences are caused by mutation alone (Shields & Helm-Bychowski, 1988). However, this also means that only allows for the matrilineal history to be studied which can lead to biased conclusions (Zhang & Hewitt, 2003) as it only reflects female mediated gene flow and female effective population size (Friesen *et al.*, 2007). Additionally, evolutionary relationships may be over simplified, genetic diversity can be underestimated and remote population processes may not be detected (Zhang & Hewitt, 2003). Clearly there are draw backs to using mtDNA alone, and it has been suggested that evolutionary relationships inferred from only a few mtDNA genes are questionable (Grealy *et al.*, 2017a) as they may not reflect the overall history (Hofreiter *et al.*, 2001; Willerslev & Cooper, 2005).

In addition to this, it has been reported that the phylogenetic relationships estimated and divergence times calculated can vary considerably based on whether the organelle or nuclear genome is used (Ksepka *et al.*, 2014; Grealy *et al.*, 2017a). For example, differences in population genetic structure were detected by different markers in humpback whales by Palumbi & Baker (1994). While the patterns between ocean and hemispheres using nuclear markers were shown to parallel that of mtDNA, when looking at a smaller spatial scale they found contrasting differences. With mtDNA markers, populations of Hawaii and California were genetically distinct but nuDNA markers showed no distinction (Palumbi and Baker, 1994). However, it is also possible for there to be a concordance between the results form mtDNA and nuDNA markers (e.g. population genetic structure in the Atlantic and Pacific Ocean in common murres (*Uria aalge*) (Morris-Pocock *et al.*, 2008).

With this in mind, we aim to use nuclear markers to investigate the population genetic structure and genetic diversity of the great auk. This will allow us to provide a more detailed picture of great auk evolution and extinction and confirm our conclusions from mitogenome data (Chapter 3).

4.2 Methods

4.2.1 Samples

Samples chosen for the nuclear marker capture were selected based on the percentage of reads retained in the preliminary mitogenome capture dataset, as a rough indication of sample

preservation and quality (Table 4.1). Samples were also chosen based on their geographical location to represent individuals from as much of the great auk's former distribution as possible (Fig. 4.1). In addition, samples from the 'Skin Mystery' (Chapter 5) were included so nuclear data would be available should the use of mitogenome data alone not resolve the mystery.

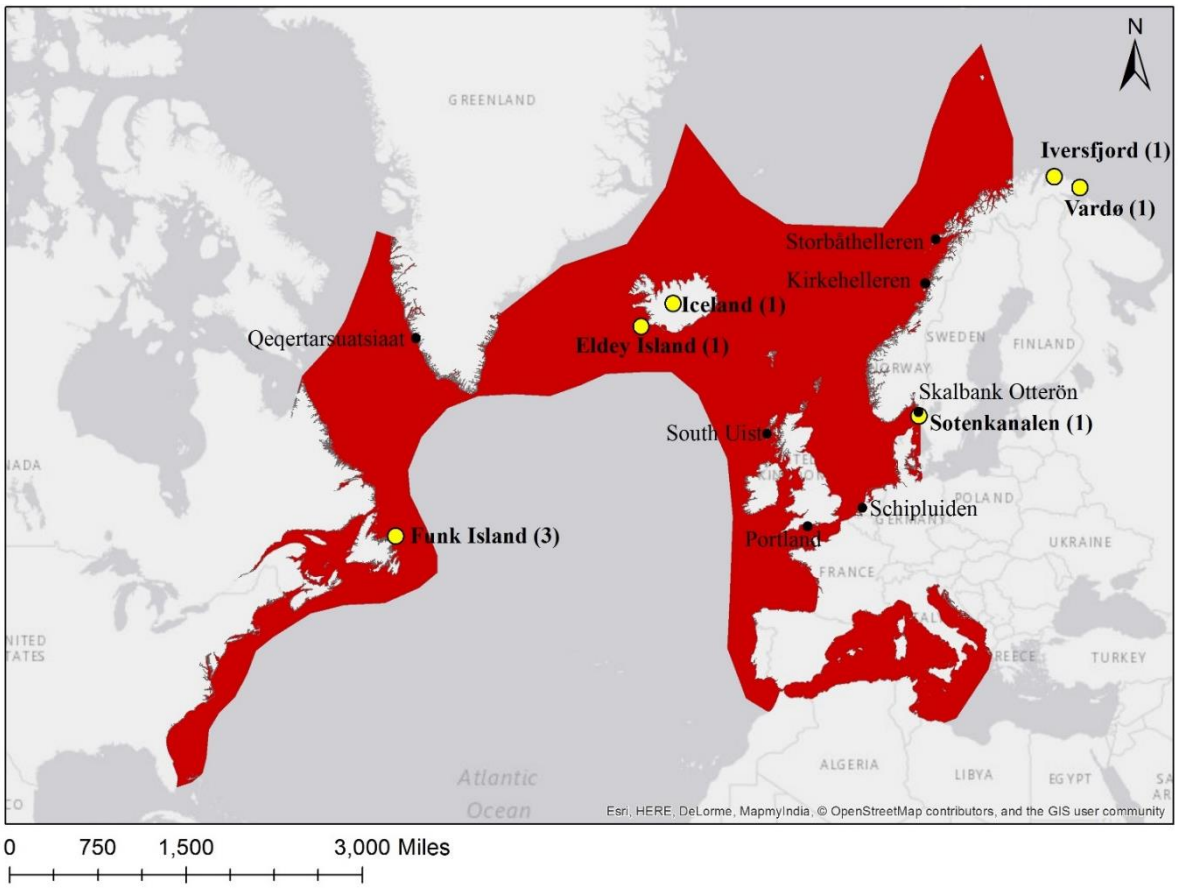

Figure 4.1 The former distribution of the great auk (red) with the locations from which material used in the nuclear population study was sourced. Yellow dots indicate samples which were sent for sequencing and numbers from that site in brackets. Black dots represent samples for which nuclear capture was performed but not sent for sequencing. Samples that fall into the 'Skin Mystery' category are not shown on map due to unknown locations, although most likely Iceland.

Sample	Location	Sample type	Approximate sample age (years before present)	% Reads retained	Post Capture Quantification (Molarity)
MK15	South Uist, Scotland	Bone	~4400-5000	0.33	0.87
MK28	Portland, England	Bone	~2000	0.32	21.85
MK48	Vardø, Norway	Bone	Unknown, unlikely >5000	1.09	7.66
MK49	Vardø, Norway	Bone	Unknown, unlikely >5000	0.44	199.73
MK50	Iceland	Bone	~200	0.36	177.25
MK51	Qeqertarsuatsiaat, Greenland	Bone	~190	0.43	37.23
MK72	Schipluiden, Netherlands	Bone	~1600-2000	2.01	2.63
MK76	Funk Island, Newfoundland	Bone	Unknown, unlikely >1000	0.46	7.33
MK78	Funk Island, Newfoundland	Bone	Unknown, unlikely >1000	0.92	649.77
MK83	Funk Island, Newfoundland	Bone	Unknown, unlikely >1000	0.26	117.77
MK91	Funk Island, Newfoundland	Bone	Unknown, unlikely >1000	1.12	18.42
MK103	Funk Island, Newfoundland	Bone	Unknown, unlikely >1000	0.43	80.37
MK104	Sotenkanalen, Sweden	Bone	~5000	0.41	13.08
MK106	Sotenkanalen, Sweden	Bone	~5000	0.91	946.36
MK108	Skalbank Otterön, Sweden	Bone	~5000	1.30	1.67
MK115	Iversfjord, Norway	Bone	~5016-4416yo	0.71	279.31
MK116	Storbåthelleren, Norway	Bone	~2100-5300	0.72	0.50
MK122	Kirkehelleren, Norway	Bone	~1500	1.07	80.10
MK130	Funk Island, Newfoundland	Bone	Unknown, unlikely >1000	0.58	48.64
MK131	Eldey Island, Iceland	Oesophagus tissue	~170	2.31	38.83
MK133	Skin mystery, probably Iceland	Toepad tissue	~170-200	0.75	136.49
MK134	Skin mystery, probably Iceland	Toepad tissue	~170-200	3.46	4.28
MK135	Skin mystery, probably Iceland	Toepad tissue	~170-200	1.23	225.03
MK136	Skin mystery, probably Iceland	Feather	~170-200	1.18	86.04

Table 4.1 Sample information (Lab ID, location, type and approximate age) for samples chosen for the nuclear marker capture. Percentage of reads retained form the mitogenome analysis was used as a rough guide to select samples of the nuclear capture. Those in yellow were selected for sequencing based on the post capture quantification results.

4.2.2 DNA extraction & library preparation

All lab work, prior to PCR amplification, was carried out in designated ancient DNA laboratories which adhered to strict ancient DNA protocols (Knapp *et al.*, 2012). For each extraction and library build, negative control samples were used to check for contamination. All post-PCR work on amplified DNA was carried out in separate laboratory facilities. Methods were chosen based on their suitability to sample type and to allow us to target the smallest fragments of DNA, typical of aDNA samples.

For bone samples, genomic DNA was extracted using Dabney et al. (2013) as this method is optimised for DNA extraction from bone (Dabney *et al.*, 2013). Genomic DNA was extracted from the oesophagus, skin, toepad tissue, and feathers using a modified version of Dabney et al. (2013) in which the initial digestion was carried out following the protocol by Gilbert et al. (2007). This digestion buffer is better suited to extraction from these tissues types than the Dabney et al. (2013) digestion buffer. Subsequent DNA purification and elution was conducted following the approach described by Dabney et al. (2013). Single stranded libraries (SSlib) were prepared for all samples following the protocol by Gansauge & Meyer (2013), with modifications as described in Bennet et al. (2014).

4.2.3 Bait design & hybridisation capture for nuclear markers

Target gene regions for hybridisation capture enrichment were selected using the following filters. Paralog genes were excluded from the capture by using UniProtIDs and EnsemblIDs in the razorbill (*Alca torda)* annotation (Gilbert, unpublished). Genes that had more than 20% of their length with missing coverage when mapping great auk reads against the razorbill genome were excluded. Great auk consensus genes were generated by replacing the razorbill genes with the homozygous SNPs found in great auk. Genes with the highest percentage divergence between the razorbill and great auk, that didn't contain any N's in their sequence, and which were less than 5kbps in length, were used to build the 20K probes resulting in 495

genes. MYcroarry probes of 120 bps long with 3X tiling (40 bps shifts) were made from CDS regions and intron regions that were adjacent to the exons of the 495 genes. Enrichment for nuclear genes was performed using MYcroarry MYbaits, following the MYcroarry Mybaits manual v3 (MYcroarray/MYbaits, 2016), using 24 hours hybridisation time, at 65ºC and final elution into 30µl nuclease free water.

Of the twenty-four samples chosen for nuclear capture, twelve were sent for sequencing (samples highlighted in Table 4.1). Those sent for sequencing were chosen using post-capture quantification results obtained from the BioAnalyser 2100. As one of our aims was to sequence the nuclear markers from the 'Skin Mystery' samples, these were included in the sequencing run, regardless of post-capture quantification. However, as three of the five skin mystery samples were already in the 'best twelve', only two samples which had better quantification results were not sent for sequencing (MK122, MK130). Therefore, in addition to the skin mystery samples (MK131, MK133, MK134, MK135, MK136), which together with MK50 represent individuals from Iceland/ assumed to be from Iceland, 6 additional samples representing individuals from Sweden (MK106), Funk Island (MK78, MK83, MK103), and Norway (MK49, MK115) were also sent for sequencing on an Illumina MiSeqPE75 platform by New Zealand Genomics Limited, Otago. Sending a subset of samples for sequencing was done so that we could initially assess whether the nuclear capture showed any evidence of success, prior to committing the full dataset.

4.2.4 Read processing

Sequencing reads were processed using the PALEOMIX v1.2.5 pipeline (Schubert *et al.,* 2014) following a procedure similar to that described by the authors. Briefly, we used AdapterRemoval v2.1.17 (Schubert *et al.,* 2016) to trim the reads for adapters and low quality bases (BaseQ < 5 or Ns), and to exclude those reads shorter than 30bp or with more than 50bp of missing data. Filtered reads from each sample were mapped against the razorbill reference genome (Gilbert, unpublished) using BWA-MEM v0.7.12 (Li, 2013), and those with low

mapping quality (MapQ < 15) removed. After the initial alignment step, Picard (v1.128, https://broadinstitute.github.io/picard) was used to exclude reads that were PCR or optical duplicates. Subsequently, GATK v3.5.0 (McKenna *et al.*, 2010) was used to perform a realignment step around indels. As we are dealing with historical samples, we also quantified the extent of DNA damage in our samples using mapDamage v2.0.6 (Jonsson *et al.*, 2013). We characterised rates of deamination in double strands (DeltaD) and single strands (DeltaS), as well as the probability of reads not terminating in overhangs (Lambda, transformed into 1/Lambda – 1, a proxy for the overhang length of overhanging regions). From these analyses, we also rescaled base quality scores according to the probability of each base being affected by post-mortem damage.

4.2.5 Calculating estimated enrichment

To determine if the capture experiments were successful, we calculated estimate enrichment levels. We first calculated the expected proportion of reads that would map if no capture was performed, by dividing the actual proportion of reads mapped to the razorbill genome by 1250 (the bait used span $1/1250^{th}$ of the genome). Subsequently the actual proportion of reads mapping to target was divided by the value for expected proportion of reads that would map if capture failed. This method was based on the calculation described by Ávila-Arcos et al. (2011).

4.3 Results

Read processing of the twelve samples initially sent for sequencing revealed low coverage of both the 495 targeted markers (0.0018x MK78 - 1.2592x MK134), and the razorbill genome overall (0.00006x MK83 - 0.0190x MK50) (Table 4.2). Full read processing stats can be found in Table S4.1a-c of the Supplementary Materials.

Sample	Country	Estimated coverage of razorbill genome	Estimated coverage of targeted genes
MK49	Norway	0.0101	0.0152
MK50	Iceland	0.0190	0.0155
MK78	Funk Island	0.0022	0.0018
MK83	Funk Island	0.00006	0.0071
MK103	Funk Island	0.0011	0.0150
MK106	Sweden	0.0172	0.0105
MK115	Norway	0.0012	0.0021
MK131	Iceland	0.0090	0.0423
MK133	Skin Mystery	0.0190	0.0154
MK134	Skin Mystery	0.0179	1.2592
MK135	Skin Mystery	0.0073	0.0106
MK136	Skin Mystery	0.0021	0.0128

Table 4.2. Estimated coverage information from the twelve samples sent for sequencing. The estimated coverage of the 495 targeted genes and estimated coverage of the reads that mapped to the razorbill genome is reported.

To determine if there were single gene regions that were sufficiently covered (coverage >= 3 to allow for majority rule consensus) in several individuals, which could be extracted and used to investigate population structure, the coverage of individual of captured genes was calculated. Table 4.3 shows the coverage range for each sample of the captured markers. We also visualised the coverage in a presence/absence matrix (Fig. 4.2). Only MK134 has any captured regions with coverage over 3, with 141 of the 495 regions having at least 3x coverage (see Fig. 4.2 & Fig. S4.1 Supplementary Materials). The coverage of individual captured markers for MK134 ranged from 0.0628 (TPK1) to 17.7232 (Ssna1). This was the only sample for which there were no markers with 0 coverage. All others had at least 78 markers with a coverage of 0, with MK78 being the worst sample with 379 of the 495 targeted genes with 0 coverage and a maximum coverage of only 0.1087. MK131 was the second most successful sample, with only 78 targeted genes with 0 coverage, however, only 1 of the genes (Ssna1) had coverage over 0.* (1.5238), so there were still no regions adequately covered for use in further analysis. This was the only other sample to have any region with over 0.* coverage for any marker. Therefore, it was not possible to work with a subset of the targeted genes and use these to investigate structure.

Sample	Country	Coverage range of captured markers
MK49	Norway	0 [125] – 0.4898 (Fam174b)
MK50	Iceland	0 [157] – 0.2204 (Isca2)
MK78	Funk Island	0 [379] – 0.1087 (Mrp130)
MK83	Funk Island	0 [223] – 0.2960 (Nipbl)
MK103	Funk Island	0 [164] – 0.7049 (Glrx5)
MK106	Sweden	0 [190] – 0.2403 (Pcp4)
MK115	Norway	0 [366] – 0.2263 (Tmem60)
MK131	Iceland	0 [78] – 1.5238 (Ssna1)
MK133	Skin mystery	0 [129] – 0.3061 (Fam174b)
MK134	Skin mystery	0.0628 (TPK1) – 17.7232 (Ssna1)
MK135	Skin mystery	0 [172] – 0.2580 (myct1)
MK136	Skin mystery	0 [142] – 0.4067 (myct1)

Table 4.3 Coverage range of captured markers. Numbers in square brackets represent the number of markers which have 0 coverage. Genes with the highest coverage are shown in brackets.

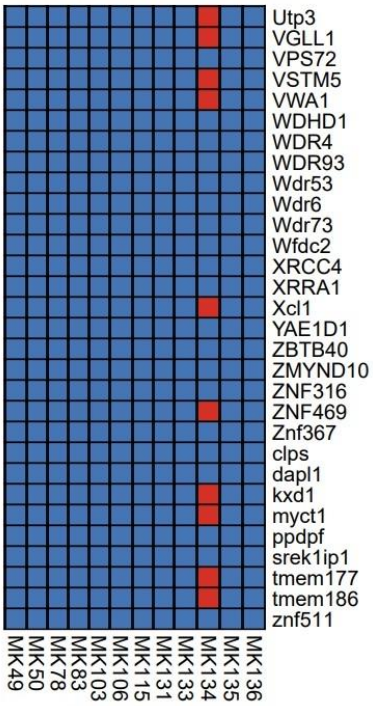

Figure 4.2 Section of the presence/absence matrix showing coverage of 30/495 captured genes (listed on the right-hand side) for each sample sent for sequencing. Presence is defined as coverage >= 3, indicated by a red square, absence is indicated by a blue square. Full presence/absence matrix can be found in the Supplementary Materials Fig. S4.1.

Estimated enrichment was calculated to confirm if the enrichment had worked (Table 4.4). Estimated enrichment levels ranged from 2.96 (MK106) – 531.16 (MK83).

Lab ID	Proportion mapping to genome (%)	Estimated proportion mapping to targets if capture failed	Proportion mapping to nuclear markers	Estimated Enrichment
MK49	0.130030271	0.000104024	0.000685235	6.59
MK50	0.197517048	0.000158014	0.000552809	3.50
MK78	0.043005673	3.44045E-05	0.00012028	3.50
MK83	0.115665162	9.25321E-05	0.04914901	531.16
MK103	0.039984493	3.19876E-05	0.001748273	54.65
MK106	0.103276174	8.26209E-05	0.00024434	2.96
MK115	0.038522099	3.08177E-05	0.00023552	7.64
MK131	0.350712129	0.00028057	0.005940535	21.17
MK133	0.199397136	0.000159518	0.000623227	3.91
MK134	0.23639317	0.000189115	0.064918447	343.28
MK135	0.135293582	0.000108235	0.000710156	6.56
MK136	0.078187077	6.25497E-05	0.001616058	25.84

Table 4.4 Estimated enrichment calculations.

4.4 Discussion

In an attempt to provide a more detailed picture of great auk evolution and extinction, in addition to the result of the mitogenome study, nuclear markers were captured and sequenced. Unfortunately, as the results showed poor coverage of the targeted nuclear markers, further analysis that would provide any meaningful results could not be performed. This discussion will focus on the possible reasons for why the results were poor and make suggestions on ways to improve the success of targeting nuDNA from the great auk samples.

4.4.1 Poor nuclear preservation or capture failure?

Target enrichment experiments often experience varying results with regards to enrichment success (Ávila-Arcos *et al.*, 2011; Enk *et al.*, 2013). Here, calculations of estimated enrichment in this experiment indicate between 2.96 (MK106) – 531.16 (MK83) fold enrichment, suggesting that the enrichment did work. It would also be unlikely that we would have 141 of the 495 nuclear markers with at least 3x coverage in sample MK134, if enrichment had failed. In MK134 enrichment levels were calculated to be 343.38fold. Interestingly, MK83, the

poorest performing sample, in terms of coverage of the razorbill genome, had the highest levels of enrichment (~530 fold). Therefore, despite the samples proving to be successful for mitogenome capture, we assume the reason for the low coverage of our targeted nuclear markers is poor DNA preservation within the samples, illustrating the points made in the introduction regarding the difficulty of sequencing nuDNA compared with mtDNA from ancient samples, and not that the capture experiments failed.

One of the biggest issues and potentially limiting factors to the field of aDNA is DNA degradation and poor DNA preservation (Pääbo, 1989; Dabney *et al.*, 2013; Damgaard *et al.*, 2015). Endogenous DNA content can vary significantly depending on source and environmental factors. For example, within sequencing libraries from bones of temperate environments there is often as little as 1% endogenous DNA (Neanderthal (Green *et al.*, 2006), with the rest being environmental contamination. Yet, in well preserved samples it can be much higher, for example, the Denisovan phalanx ~70% (Reich *et al.*, 2010), which is closer to permafrost preserved samples than temperate environments (Poinar *et al.*, 2006; Rasmussen *et al.*, 2010). Endogenous content of samples used in this experiment was not calculated, however, the samples sent for sequencing ranged in age from approximately 170 - 5000 years old and are all from the more northerly locations of the great auks' former range. This is not particularly old for aDNA samples, with the oldest whole genome sequenced being ~735,000 years old (Orlando *et al.*, 2013), even compared with samples used in our mitogenome capture (oldest ~12,000-15,000 years old from a cave in northern Spain) and as they come from areas where the climate was colder and potentially more stable, conditions were likely to be better for DNA preservation (Smith *et al.*, 2001).

Different sources of aDNA, e.g. bone, teeth, tissue, and egg shell, are known to have varying levels of preservation, with regards to the endogenous content. For example, a comparison experiment between bone and teeth showed an average endogenous content of 2.2% for parietal skull bone, 16.4% in teeth and 40% for petrous bones (Hansen *et al.*, 2017)). Egg shell has

demonstrated excellent levels of preservation, with the retrieval of both mtDNA and nuDNA, with moa eggshell also showing 125 times lower bacterial load than bone of same age (Oskam *et al.*, 2010). Within this study, our sources of aDNA were bone, feather and tissue from the toe pad, body and oesophagus, which was stored in spirit. Sample MK134, the sample with the highest coverage, was extracted from a toepad tissue sample from a specimen thought to be ~170-200 years old, most likely from Iceland. Toepad tissue has been used in several aDNA studies with great success, for example, the first nuclear DNA to be obtained from the extinct passenger pigeon (*Ectopistes migratorius*) (Fulton *et al.*, 2012) was extracted from a toepad sample, and since then toepads have been used to sequence whole genomes from the passenger pigeon and used in demographic reconstructions (Hung *et al.*, 2014; Murray *et al.*, 2017). The calamus (quill) of feathers has been used as a source of DNA for decades (Taberlet & Bouvet, 1991; Leeton *et al.*, 1993; Payne & Sorenson, 2003), and more recently studies have shown that it is possible to successfully extract aDNA from the rachis (main shaft) and barbs (Rawlence *et al.*, 2009). One feather sample was used in this study, MK136, however it did not perform better than the bone or tissue samples used. MK131, the second-best sample, is an oesophagus tissue sample from one of the last individuals killed in June 1844, from Eldey Island, Iceland. This sample has been stored in spirit (Newton, 1861), thought to originally be whisky, but has since been 'topped up' with ethanol. Therefore, the ethanol content is likely to be high, allowing for good DNA preservation. Bone samples are commonly used in aDNA studies, as one of the components of a vertebrates body likely to survive a long time (Green & Speller, 2017). In avian aDNA studies, Grealy et al. (2017a) report that half of studies used DNA extracted from single-source bone. In this study, bones form the bulk of samples tested (7/12 sent for sequencing were bone samples). The worst performing sample, in terms of coverage of targeted markers, was MK78. This bone sample is from Funk Island, and likely to be one of the youngest (most likely less than 500 years old). While Funk Island is from a high latitude location, found in the Low Arctic oceanographic region, many of the bones excavated from Funk Island come from the surface layers. This means they have potentially been exposed to changes in environmental conditions, e.g. temperature and moisture, and also exposed to UV

light damage from the sun. As Funk Island is still home to thousands of birds, there is also the consideration that the substrate the bones are in is highly acidic due to it consisting of mainly guano, built up over thousands of years. All these factors combined could lead to poorly preserved DNA. In this thesis, bone samples were chosen preferentially over other sample sources as they had provenance details required for the population genetic study. However, as it has been noted that bones can contain very low proportions of endogenous DNA and the quantity and quality of DNA extracted from them can vary greatly (Green & Speller, 2017), which is evident in our nuclear and mitochondrial studies of the great auk, perhaps other sources may have given a greater success rate in the nuclear capture.

Interestingly, the two samples that exhibited greatest enrichment, MK134 and MK131, in terms of highest average coverage and the only samples to have at least one marker covered >1-fold coverage, were those which had the highest percentage of reads retained during the mitogenome capture. However, their post-capture quantification results were comparatively low, compared to many of the other samples, and they were only selected for this first run of sequencing due to being from the 'Skin Mystery'. Therefore, perhaps post-capture quantification does not indicate success for capture, but more likely those samples with the highest percentage of reads retained should have been sent for sequencing in spite of low levels of post-capture quantification as this could simply be reflective of non-endogenous DNA, despite using targeted enrichment for the great auk nuclear markers. This is also reflected by samples which had the highest post-capture quantification results (MK106, MK78, MK115) but some of the lowest average coverage of targeted markers.

4.4.2 How could we improve this study?

Since the first aDNA experiments of the 1980s, protocols are continually being developed to improve the success of aDNA extraction, amplification and sequencing. While the methods employed in this study were suitable for the sample type, and the degraded nature of the aDNA, since this project began and the lab work completed, a number of protocols have been published

which potentially could have increased our chances of sequencing nuclear markers from our great auk samples. Additionally, there are a number of variables that have been suggested that affect capture efficiency: initial aDNA input, hybridisation time and bait length distribution (Ávila-Arcos *et al.*, 2015), which should be considered for future attempts of nuclear capture from the great auk samples.

Several studies have reported that capture efficiency is limited by the degree of complexity in the initial aDNA library or by clonality (Ávila-Arcos *et al.*, 2011, 2015; Carpenter *et al.*, 2013). Clonality in our samples ranged from 1-83% (Supplementary Materials Table S4.1a-c). Carpenter et al. (2013) suggested implementing the SSlib build method Gansauge & Meyer (2013) to increase complexity by retaining more of the smaller fragments lost during other library preparation methods. However, this method was used in this study. Recently, a single-tube library preparation method, for degraded DNA has been published which states it increases library complexity, and yields more reads that map uniquely to the reference genome, in addition to being cheaper and quicker that previously published methods (Carøe *et al.*, 2017). Therefore, it is possible that using this library preparation method would have given us a greater chance of success with our great auk samples.

The endogenous content of a sample is also important for successful capture (Ávila-Arcos *et al.*, 2011, 2015), therefore, success may have been improved if samples used in the capture had the highest endogenous content. Damgaard et al. (2015) have shown that applying a 'pre-digestion' step in DNA extraction from bone, increased the proportion of endogenous DNA. However, they also state that the pre-digestion step is not recommendable when only a small amount of starting material is available (<50mg), as it could in fact cause the final DNA concentration in the extract to be critically low (Damgaard *et al.*, 2015). With regards to the great auk bones sampled for this study, the amount used in extraction was on average 46.73mg. Therefore, despite the proven success of the pre-digestion step in increasing endogenous

content, for our samples it would not have been advisable to use, unless we were able to resample the bones and collect more bone powder.

Additionally, the manufacturers of the bait used in this study (MYbaits) suggest three modifications that can be made to the protocol for use with aDNA samples. Firstly, modifying the hybridisation and wash temperature. In this study, we used a temperature of 65^{o}C. MYbaits suggest a temperature of 55^{o}C can improve the captured target complexity, whereas 60-65^{o}C can give a higher on-target percentage (MYcroarray, 2016). Therefore, perhaps the optimal temperature was not used, and we should have performed trials to find the best temperature for our bait and libraries. Secondly, they suggest modifying the hybridisation time. In our protocol, we used a hybridisation time of at least 24 hours, but MYbaits suggest using between 24-40 hours, again using trials to optimise (MYcroarray, 2016). Ávila-Arcos et al. (2015) commented on the effect of hybridisation times, suggesting that it was one of the variables which may affect capture efficiency, however, they did not directly compare hybridisation times within the same protocols, only between two different whole-genome capture methods. The final suggestion by MYbaits is to perform a second round of capture to improve specificity (raw reads mapping to targets) (MYcroarray, 2016). However, this can also increase clonality which of course is another issue.

Another consideration is regarding how samples were selected for capture. In this study, samples were selected based on post-capture quantification results, however, as shown by our results this was not a good indication for capture success. Here, the better indicator was the percentage of reads retained in the mitogenome capture (as evident from sample MK134). Another method that could have been used as a predictor of enrichment success is quantitative PCR (qPCR) (Enk *et al.*, 2013). However, while this method may be appropriate for some target enrichment experiments (such as mitochondrial regions or a smaller number of targets), we feel that due to the complexity of sequencing 495 markers, it would not have been appropriate. This is evident from the results of MK134. Here we can see that not all markers are evenly covered. Therefore, depending on which markers were chosen to be screened for

using qPCR, assuming a select few were chosen, the results of said chosen marker may be one which is poorly covered, but others in the sample are good but not screened for. Therefore, we could disregard a sample if the marker did not produce a good qPCR result and therefore not send potentially some of the best samples for sequencing. Additionally, as the fragment size of the aDNA reads is very low, especially as the Dabney et al. (2013) methods, increases the number of the smaller fragments successfully extracted, the size restriction of qPCR assays (generally requiring templates of at least 60bp in length) would likely render the method uninformative.

Ávila-Arcos et al. (2015) suggest initial shotgun screening of samples, to obtain information regarding endogenous content, length distribution and library complexity, is crucial for deciding if a sample is suitable for whole-genome capture (WGC). While we were not performing WGC in this study, this information would have been useful in determining which samples may have yielded the best results following capture of nuclear markers and potentially allowed us to obtain useable data for population genetic analysis.

4.5 Conclusion

Despite the results of this nuclear marker study proving too low to perform meaningful analysis, it has revealed insights into the preservation of nuclear DNA in great auk material. We have concluded that the low coverage was due to poor sample preservation and not that the capture experiment had failed. We also suggest that in future, if initial shotgun sequencing to obtain information of endogenous content or library complexity cannot be performed, then the percentage of reads retained from mitogenome work is a better indicator for capture success than the post-capture quantification results.

4.6 Additional information

Author contributions

J.E.T., G.R.C., M.T.P.G., and M.K. conceived the study and designed the experiments; J.A.S.C. designed the capture bait; J.E.T. performed the experiments; F.J.G.V. and J.E.T analysed the data; N.J.R., M.H., and J.R.S. provided the initial framework for the overall great auk population study; all authors contributed to writing the paper.

Chapter 4: Supplementary Materials

S4 Tables

Table S4.1a-c Summary data for nuclear capture.

Lab ID	Total number of pairs	Total number of retained reads	Average number of NTs in retained reads
MK49	2155408	2100011	56.33
MK50	2512005	2518051	58.71
MK78	1889190	1488199	52.25
MK83	11576	11516	62.96
MK103	716586	673236	56.09
MK106	4944184	4551040	55.81
MK115	976779	823709	49.76
MK131	1017918	775351	47.86
MK133	2808065	2838451	62.84
MK134	1616879	1548928	60.96
MK135	1559426	1527833	59.08
MK136	859077	803189	58.30

Table S4.1a Summary data for samples used in nuclear capture. Details on total number of pairs and reads retained.

Lab ID	Total number of hits (prior to PCR duplicate filtering)	Total number of hits vs. total number of reads retained	Fraction of hits that were PCR duplicates	Total number of hits (excluding PCR duplicates)	Total number of unique hits vs. total number of reads retained	Estimated coverage from unique hits	Average number of aligned bases per unique hit
MK49	305282	0.1454	0.1055	273065	0.1300	0.01011	43.24
MK50	624862	0.2482	0.2041	497358	0.1975	0.01899	44.61
MK78	68919	0.0463	0.0714	64001	0.0430	0.00222	40.47
MK83	2527	0.2194	0.4729	1332	0.1157	0.00007	58.19
MK103	29276	0.0435	0.0805	26919	0.0400	0.00109	47.18
MK106	511437	0.1124	0.0810	470014	0.1033	0.01720	42.77
MK115	35866	0.0435	0.1153	31731	0.0385	0.00123	45.34
MK131	287908	0.3713	0.0555	271925	0.3507	0.00896	38.50
MK133	634070	0.2234	0.1074	565979	0.1994	0.01895	39.13
MK134	952897	0.6152	0.6157	366156	0.2364	0.01795	57.27
MK135	234629	0.1536	0.1190	206706	0.1353	0.00730	41.25
MK136	74151	0.0923	0.1531	62799	0.0782	0.00213	39.56

Table S4.1b Summary data for samples used in nuclear capture when mapped to razorbill genome.

Lab ID	Total number of hits (prior to PCR duplicate filtering	Total number of hits vs. total number of reads retained	Fraction of hits that were PCR duplicates	Total number of hits (excluding PCR duplicates	Total number of unique hits vs. total number of reads retained	Estimated coverage from unique hits	Average number of aligned bases per unique hit
MK49	2301	0.0011	0.3746	1439	0.0007	0.0152	51.52
MK50	8359	0.0033	0.8335	1392	0.0006	0.0155	54.22
MK78	185	0.0001	0.0324	179	0.0001	0.0018	48.73
MK83	1638	0.1422	0.6545	566	0.0491	0.0071	61.02
MK103	1479	0.0022	0.2042	1177	0.0017	0.0150	62.13
MK106	1190	0.0003	0.0655	1112	0.0002	0.0105	45.89
MK115	196	0.0002	0.0102	194	0.0002	0.0021	54.02
MK131	6219	0.0080	0.2594	4606	0.0059	0.0423	44.83
MK133	3908	0.0014	0.5473	1769	0.0006	0.0154	42.38
MK134	522228	0.3372	0.8075	100554	0.0649	1.2592	61.12
MK135	1257	0.0008	0.1368	1085	0.0007	0.0106	47.65
MK136	1919	0.0024	0.3236	1298	0.0016	0.0128	48.05

Table S4.1c Summary data for samples used in nuclear capture when mapped to targeted markers.

S4 Figures

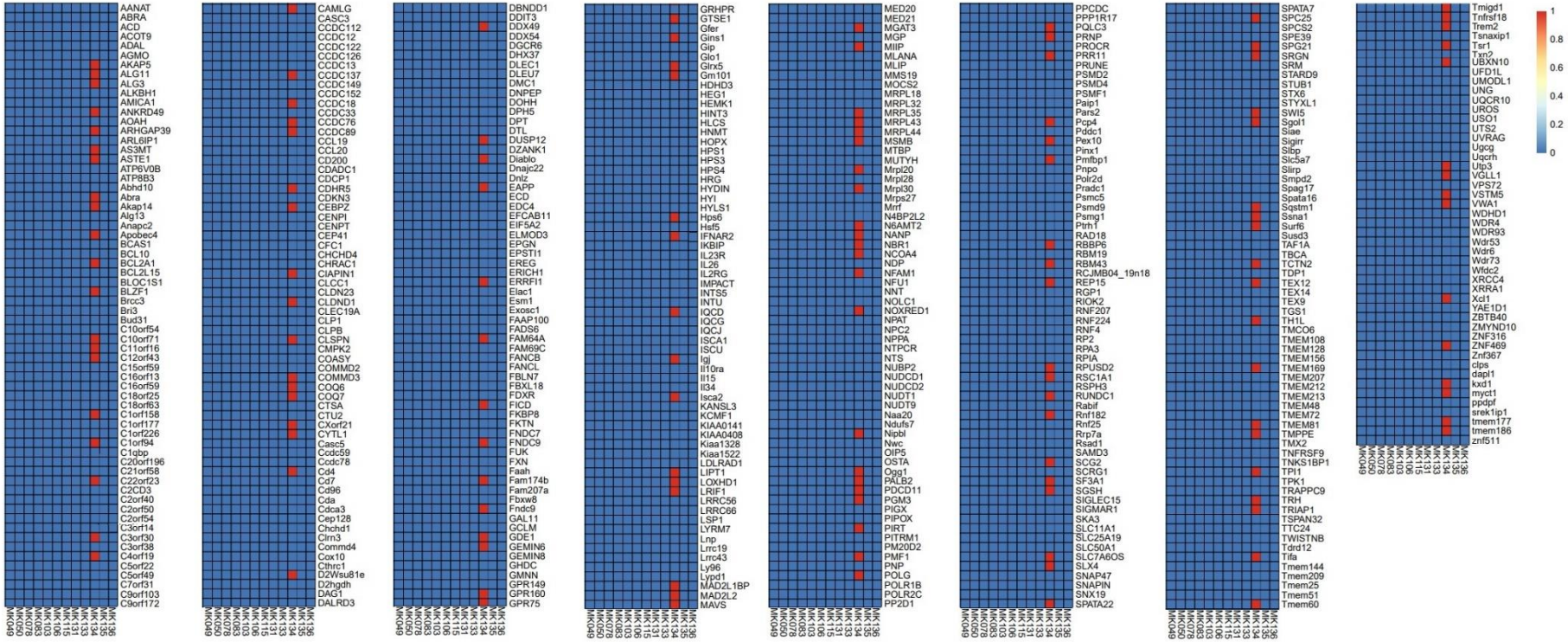

Figure S4.1 Full presence/absence matrix of coverage of nuclear markers.

Chapter 5: An ʻAukwardʼ Tale: A Genetic Approach to Discover the Whereabouts of the Last Great Auks[3]

Jessica E. Thomas, Gary R. Carvalho†, James Haile†, Michael D. Martin, Jose A. Samaniego Castruita, Jonas Niemann, Mikkel-Holger S. Sinding, Marcela Sandoval-Velasco, Nicolas J. Rawlence, Errol Fuller, Jon Fjeldså, Michael Hofreiter, John R. Stewart, M. Thomas P. Gilbert† and Michael Knapp†(Equal contribution †)

[3] This chapter has been published. Thomas J. E., Carvalho G. R., Haile, J., *et al*., (2017) An ʻAukwardʼ Tale: A Genetic Approach to Discover the Whereabouts of the Last great auks. *Genes* **2017**, *8*(6), 164; doi:10.3390/genes8060164. Appendix 2.

Abstract

One hundred and seventy-three years ago, the last two great auks, *Pinguinus impennis*, ever reliably seen were killed. Their internal organs can be found in the collections of the Natural History Museum of Denmark, but the location of their skins has remained a mystery. In 1999, great auk expert Errol Fuller proposed a list of five potential candidate skins in museums around the world. Here we take a palaeogenomic approach to test which—if any—of Fuller's candidate skins likely belong to either of the two birds. Using mitochondrial genomes from the five candidate birds (housed in museums in Bremen, Brussels, Kiel, Los Angeles, and Oldenburg) and the organs of the last two known individuals, we partially solve the mystery that has been on great auk scholars' minds for generations and make new suggestions as to the whereabouts of the still-missing skin from these two birds.

5.1 Introduction

Over the past three decades, the field of ancient DNA (aDNA) has grown considerably, from sequencing a small section of mitochondrial DNA from the Quagga, an extinct form of the plains zebra (Higuchi *et al.*, 1984), to whole genome sequencing from samples up to 735,000 years old (Orlando *et al.*, 2013). Ancient DNA has been used to answer and address a diverse range of ecological and evolutionary questions, providing insight into countless species' pasts, including our own. However, aDNA can also be a useful tool for museums, specifically for species identification and, under suitable circumstances for reconstructing the history of specimens where museum records are insufficient. This study traces the whereabouts of the skins from the last two documented great auks using a palaeogenomic approach.

The great auk (Fig. 5.1), *Pinguinus impennis*, Bonnaterre (1790) (traditionally *Alca impennis*, Linnaeus, 1758), has been described as "*perhaps the most curious of all vanished birds*" (Fuller, 1999). It was a bird whose life and ultimate extinction has generated ongoing interest, with several scholars dedicating their lives to great auk research (Newton, 1861; Grieve, 1885; Bengtson, 1984; Fuller, 1999; Gaskell, 2000). Even now, 173 years after the death of the last two recorded captured individuals, there are still many unanswered questions concerning aspect of its life-history, evolution, and extinction.

Figure 5.1 A mounted great auk skin, The Brussels Auk (RBINS 5355) (MK135), from the collections at Royal Belgian Institute of Natural Sciences (Credit Thierry Hubin (RBINS)).

One such mystery that surrounds the great auk is the whereabouts of the skins from the last documented pair. In order to be able to correlate the phenotype of the last birds with genomic information obtained from the well-preserved organs, and in view of the active role that researchers and research institutions played in pushing the great auk towards extinction, it is of relevance to be able to trace these skins.

Once found in great numbers across the North Atlantic (Fig. 5.2), this flightless bird was heavily hunted for its meat, oil, and feathers. By the start of the 19th century, populations in the North-West Atlantic had been decimated. The last few remaining birds were breeding on the skerries off the south-west coast of Iceland, but with their scarcity increasing, great auks were then also sought after as a desirable item for both private and institutional collections (Bengtson, 1984; Meldgaard, 1988; Montevecchi & Kirk, 1996; Fuller, 1999; Serjeantson, 2001).

Figure 5.2 The great auks' former range across the North Atlantic, as indicated by the red area and the location of Eldey Island (yellow dot) off the south-west coast of Iceland, the site where the last documented great auks were killed. Maps were created using spatial data provided by BirdLife International/IUCN (BirdLife International & IUCN, 2016) with the National Geographic basemap in ArcGIS (ESRI, Redlands, CA, USA) (ESRI, 2016) .

From 1830 to 1841, several trips were taken to Eldey Island (Fig. 5.2) where great auks were caught, killed, and sold for exhibitions. Following a three-year period of no recorded captures of great auks, Carl Siemsen commissioned an expedition to Eldey to search for any remaining birds. Between 2 and 5 June 1844, the expedition reached Eldey Island where two great auks were observed amongst smaller birds inhabiting the island. Both Auks were killed and their broken egg discarded. The birds, though, were never to reach Siemsen. The expedition leader sold them to Christian Hansen, who then sold them to the apothecary Möller, in Reykjavik, Iceland. Möller skinned the birds and sent them, as well as their preserved body parts, to Denmark (Newton, 1861; Fuller, 1999; Gaskell, 2000).

The internal organs of these two birds now reside in the Natural History Museum of Denmark. However, the location of the skins of those individuals remains a mystery, despite considerable effort of notable great auk scholars to solve it. Fuller (1999) describes in detail the known history of the 80 or so specimens that are still in existence in collections today and concludes: "*Somehow, amid all the frantic Garefowl* [another name for great auk] *research of the nineteenth century, they* [the skins] *were lost track of. Several of the surviving stuffed specimens, notably those in Kiel, Bremen and Oldenburg were tentatively identified with them. The most likely candidates, however, are the birds now in Los Angeles and in Brussels*" (Fuller, 1999) (p. 85).

Our study compares complete mitochondrial genome (mitogenome) sequences from the five candidate skins (those housed in Bremen, Brussels, Kiel, Los Angeles, and Oldenburg) to the internal organs of the last documented captured great auks (stored in Copenhagen) to test which-if any-of Fuller's candidate skins likely belong to one of the last two individuals.

5.2 Methods

5.2.1 Sample information

Specimens from the candidate list proposed by Fuller (1999) and the organs from the two 1844 Eldey Island individuals, were sampled using sterile equipment and the appropriate method for sample type, which caused minimal physical damage to the specimen (Table 5.1).

Lab ID	Bird Name & Description	Origin & Date	Institution	Curator/ Collector	Institution Number	Sample Type/ Method
MK131	Last great auk 1. Oesophagus (male).	Eldey Island, Iceland. Date: June 1844	Natural History Museum of Denmark. Copenhagen, Denmark	J.Fjeldså/ J.Thomas	NHMD 153069	Oesophagus tissue cut from the end.
MK132	Last great auk 2. Oesophagus (female).	Eldey Island, Iceland. Date: June 1844	Natural History Museum of Denmark. Copenhagen, Denmark	J.Fjeldså/ J.Thomas	NHMD 153070	Oesophagus tissue cut from the end.
MK133	The Oldenburg Auk. Adult in summer plumage.	Iceland. Probably Eldey. Date: Unknown	Landesmuseum Natur und Mensch Oldenburg. Germany	C.Barilaro	AVE 8086	Body tissue cut from under wing.
MK134	The Bremen Auk. Adult in summer plumage.	Unknown, probably Eldey. Date: Unknown	Übersee-Museum Bremen. Germany	M.Stiller	RKNr. 2392	Toepad tissue cut from feet.
MK135	The Brussels Auk. Adult in summer plumage.	Probably Eldey. Date: Unknown perhaps June, 1844	Institut Royal des Sciences Naturelles de Belgique. Brussels, Belgium	G.Lenglet	RBINS 5355	Toepad tissue. Tissue cut from feet
MK136	Dawson Rowley's Los Angeles Auk. Adult in summer plumage, possibly female.	Iceland, probably Eldey. Date: Unknown perhaps June, 1844	Natural History Museum of Los Angeles County. USA	K.Garett	LACM 76476	Feather plucked from body.
MK138	The Schleswig-Holstein Auk. Adult in summer plumage.	Unknown. Date: Unknown	Zoologisches Museum der Christian-Albrechts Universität zu Kiel. Germany	D.Brandis/ L.Rosotta	cat. No. A0585	Toepad tissue cut from feet.
LastGA2_Heart	Last great auk 2. Heart (female).	Eldey Island, Iceland. Date: June 1844	Natural History Museum of Denmark. Copenhagen, Denmark	J.Fjeldså/ J.Haile	NHMD 153070	Heart tissue cut from aorta.

Table 5.1 Sample information. Lab ID number used during laboratory and analysis process. Mount name and description given by Fuller (1999). Origin and date information as noted by Fuller (1999). Institution information relating to the present location of specimen and the curator/sample collector name.

5.2.2 DNA extraction

All lab work prior to polymerase chain reaction (PCR) amplification was carried out in designated aDNA laboratories that adhere to strict aDNA protocols (Knapp *et al.*, 2012). For each DNA extraction and library build, negative controls were used to check for contamination by exogenous DNA. All post-PCR work on amplified DNA was carried out in separate laboratory facilities.

Genomic DNA was extracted from the oesophagus (Fig. 5.3a), skin (Fig. 5.3b), toepad tissue (Fig. 5.3c), and feathers using a modified version of Dabney et al. (2013) in which the initial digestion was carried out following the protocol by Gilbert et al. (2007). This digestion buffer is better suited to extraction from these tissues types than the Dabney et al. (2013) digestion buffer, which was optimised for DNA extraction from bone. Subsequent DNA purification and elution was conducted following the approach described by Dabney et al. (2013). Genomic DNA was extracted from the heart tissue (Fig. 5.3d) using the protocol by Campos & Gilbert (2012).

(a) **(b)** **(c)** **(d)**

Figure 5.3 (**a**) Jars containing the oesophagus from the last two individuals killed on Eldey Island (NHMD153069/NHMD153070) (Credit: J. Thomas). (**b**) Sampling of The Oldenburg Auk (AVE 8086) to remove a section of body tissue for DNA extraction (Credit: C. Barilaro, Landesmuseum Natur und Mensch Oldenburg). (**c**) Sampling the toe pad of The Bremen Auk (RKNr. 2392) to remove tissue sample (Credit: M. Stiller, Übersee-Museum Bremen). (**d**) The hearts from the last two documented individuals. The heart from the female individual has been sampled for this study (top) (NHMD153070) (Credit: Natural History Museum of Denmark).

5.2.3 Data generation

Single stranded libraries were constructed for all samples, except LastGA2_Heart, following Gansauge & Meyer (2013), with modifications as described by Bennett et al. (2014), as this allowed for targeting of the smallest fragments of DNA, typical of highly degraded specimens. For LastGA2_Heart, the protocol described by Meyer & Kircher (2010) was used. Enrichment for complete mitogenomes was performed using MYcroarray MYbaits, following the manufacturer's manual v2.3.1 (MYcroarray MYbaits, 2014) on all samples except MK138 and LastGA2_Heart. Samples were sequenced on Illumina platforms (HiSeq and MiSeq) by New Zealand Genomics Limited, Otago Branch, or the Danish National High-Throughput DNA Sequencing Centre.

5.2.4 Read processing

Processing of raw sequence data was facilitated by the PALEOMIX v1.2.5 pipeline (Schubert *et al.*, 2014), which performs adapter trimming, read mapping to a reference genome, and quality-based filtering. Low-quality bases and adapter sequences were trimmed from the 3' ends of DNA reads with the software AdapterRemoval v2.1.7 (Lindgreen, 2012; Schubert *et al*., 2016) using a mismatch rate of 0.333 (command-line option—mm 3). Paired end reads overlapping by at least 11 base pairs (bp) were collapsed into a single read with re-calibrated base quality scores. Trimmed reads shorter than 25 bp were discarded.

Mapping to the great auk reference mitogenome (GenBank: KU158188.1) (Anmarkrud & Lifjeld, 2017) was performed with Burrows–Wheeler Aligner (BWA) v0.5.10 (Li & Durbin, 2009) with seeding deactivated and otherwise default settings. PCR duplicates were removed with the MarkDuplicates function within Picard v1.82 (Broad Institute) and the rmdup function within the software SAMtools (Li *et al.*, 2009). Collapsed reads were filtered using a script included with PALEOMIX. Reads with mapping quality (MAPQ) scores <20 were removed from further analysis. Local realignment of reads misaligned to the reference mitogenome was performed with the RealignerTargetCreator and IndelRealigner tools included in the software Genome Analysis Toolkit (GATK) v3.6.0 (McKenna *et al.*, 2010). The pipeline also utilised MapDamage2 (Jonsson *et al.*, 2013) to recalibrate base qualities of aligned sequence reads in each sequencing library in order to remove the residual aDNA damage patterns. The

UnifiedGenotyper algorithm within GATK v3.6.0 was used to determine haploid genotypes within individual samples.

A relaxed and strict filtering system was used to create consensus sequences and alignments from the processed data. In the first stage of filtering, both systems used VCFtools (Danecek *et al.*, 2011) to filter genotypes from the final alignment when their genotype quality scores were less than 30. For the relaxed alignment, the per-individual read depth was set to only include bases with a minimum of 3-fold coverage. Bases called for the consensus sequence had to be present at a frequency higher than 33%. To be included in the final alignment, no more than 33% of bases could be missing from the consensus sequence of an individual.

For the strict settings, the per-individual read depth was set to only include bases with at least 10-fold coverage. Geneious v-10.1.3 (Kearse *et al.*, 2012) was used to filter bases so that the majority base was present in more than 90% of reads. For an individual to be included in the final alignment, no more than 20% of sites could be missing from the individual's consensus sequence. A custom script was used to convert the filtered Variant Call Format (VCF) file into a multiple sequence alignment in FASTA format.

Following read processing, the data was aligned using Seaview v4.0 (Gouy, Guindon and Gascuel, 2010) with the algorithm *Muscle -maxiters2 -diags*. The alignment was manually checked for errors using BioEdit v7.2.5 (Hall, 1999), and Tablet v-1.16.09.06 (Milne *et al.*, 2013), was used to view the rescaled Binary Alignment Map (BAM) file for each sample.

MEGA v-7.0.21 (Kumar *et al.*, 2016) was used to generate a pairwise distance table for all sequenced individuals. Phylogenetic relationships between the individuals were reconstructed and visualised using a maximum-likelihood approach as implemented in MEGA v-7.0.21 (Kumar *et al.*, 2016). jModelTest v-2.1.10 (Guindon and Gascuel, 2003; Darriba *et al.*, 2012) was used to determine the most suitable nucleotide substitution model, which was a Hasegawa–Kishino–Yano (HKY) (Hasegawa *et al.*, 1985) model. Initial trees for the heuristic search were obtained by applying Neighbour-Joining methods to a matrix of pairwise distances estimated using the maximum composite likelihood approach. Branch lengths are measured in number of substitutions per site. All positions containing gaps and missing data were removed. Phylogenies were reconstructed from 500 bootstrap pseudo replicates to evaluate branch support.

5.3 Results

Mitogenome sequence data was obtained from all candidate specimens as well as from the two oesophagi of the last great auks. Unique coverage of the mitogenomes for these samples ranged from 6.2× to 288.6× (Table 5.2). As DNA extracted from the oesophagus of the female last great auk (MK132) yielded only a low coverage, poor quality mitogenome assembly, DNA from the heart of the same individual was also sequenced. This yielded a high coverage (430×) mitogenome, which was used in all further analyses.

Sample	GenBank Accession Number	Number of Reads	Number of Unique Reads Mapping to Ref. Mitogenome	Estimated Coverage	Relaxed Settings Seq. Length (bp [1])	Strict Settings Seq. Length (bp)
MK131	MF188883	300754 (read pairs)	30,297	74.40	16,001	15,067
MK132	NA	550631 (read pairs)	2366	6.23	13,267	3312
MK133	MF188884	429392 (read pairs)	8750	23.04	16,251	14,240
MK134	MF188885	343766 (read pairs)	86,325	288.62	16,607	16,526
MK135	MF188886	579992 (read pairs)	27,767	88.90	16,554	16,356
MK136	MF188887	563635 (read pairs)	24,401	67.83	16,330	15,833
MK138	MF188888	10796460 (SE[2] reads)	2799	9.76	16509	7866
LastGA2 Heart	MF188889	957970612 (SE reads)	121,886	430.09	16,698	16,649

[1] Base pairs (bp); [2] Single End (SE).

Table 5.2 Read processing results for all samples. Information displayed: GenBank Accession number, the number of reads (PE/SR), number of unique reads mapping to the reference mitogenome, estimated coverage from unique reads and sequence length using the relaxed and strict settings.

With the sequence data from the heart of the female last great auk (LastGA2_Heart), the alignment of all sequences assembled under the relaxed rules had a length of 15,790 bp after sites not covered by all consensus sequences were removed. For the strict alignment, MK138 did not meet criteria set by the strict filtering settings as more than 20% sites were missing. With this individual removed, we obtained a strict alignment length of 13,475 bp. The pairwise distance matrix (Table 5.3) shows that the consensus sequence obtained from sample MK131, the oesophagus of the male, is identical to the consensus sequence obtained from MK135, The

Brussels Auk. No other consensus sequences match. LastGA2_Heart, the female last great auk, groups with MK136 and MK134 in the maximum likelihood phylogeny (Fig. 5.4), but there are 18 and 20 well-supported differences between the consensus sequences, respectively. Analysis presented here was generated using data from the relaxed filtering settings, but results were consistent with data from the strict filtering system. Thus, only the male last great auk has a corresponding DNA match among the candidate skin samples identified by Fuller (1999).

	MK131	MK133	MK134	MK135	MK136	MK138	LastGA2_Heart
MK131_LastGA1							
MK133_Oldenburg	17						
MK134_Bremen	18	23					
MK135_Brussels	0	17	18				
MK136_LA	16	23	20	16			
MK138_Kiel	14	11	20	14	20		
LastGA2_Heart	16	23	20	16	18	20	

Table 5.3 Pairwise distance matrix. Estimates of evolutionary divergence between sequences generated using the relaxed settings. The number of base differences per sequence from between sequences are shown. All positions containing gaps and missing data were removed, leaving a total of 15,790 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar *et al.*, 2016).

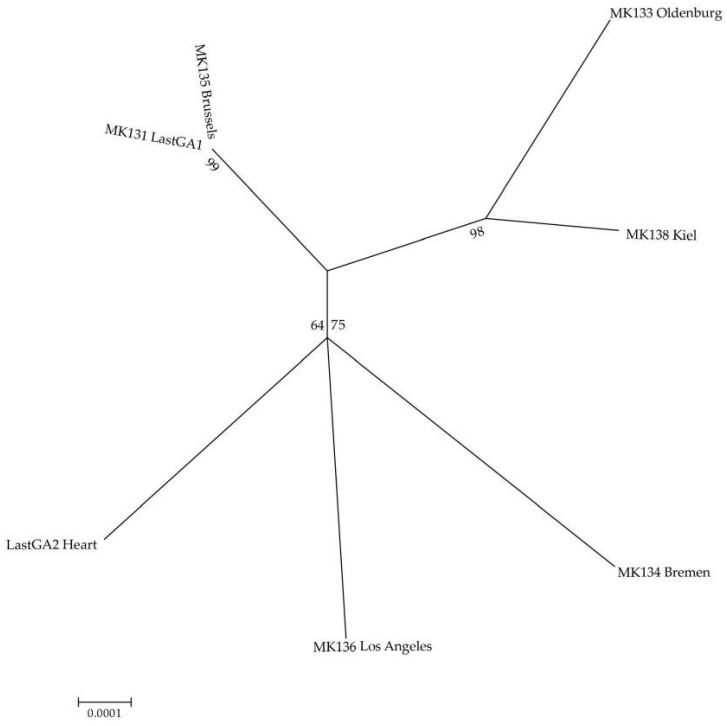

Figure 5.4 Maximum likelihood reconstruction of phylogenetic relationships between individuals, under the relaxed filtering settings. Branch labels are bootstrap support values for the respective sample. Evolutionary analyses were conducted in MEGA7 (Kumar *et al.*, 2016).

5.4 Discussion

The genetic analyses presented here help to partially resolve the mystery of the missing skins of the last two great auks. They provide evidence of matching mitochondrial genomes for the internal organs of the last male great auk held at the Natural History Museum of Denmark in Copenhagen and the great auk skin held at the Royal Belgian Institute of Natural Sciences, Brussels (Fig. 5.1). Mitochondrial DNA cannot always be unambiguously used in identification of individuals. However, in a broader study of forty-one great auk mitogenomes from across their range, Thomas et al. (in prep), found that mitochondrial diversity in great auks remained high right up to their demise, with no other individuals found to have the same mitochondrial haplotype. Together with the information from the historical record, the match between the internal organs and The Brussels Auk therefore appears to be more than just a coincidence.

There are around 80 known mounted great auk skins in museums worldwide. However, the majority can be ruled out of any speculation that they belonged to the last pair due to their history (for example, if they were collected before 1844). Those tested in the current study were placed on the candidate specimen list due to several factors that led Fuller, as well as other experts like the University of Copenhagen Professor Japetus Steenstrup (dubbed 'Father of Garefowl History' by Grieve, 1885), and Grieve (1885), to suspect that they originated from the 1844 Eldey pair. Details such as when and where they were acquired, from whom (i.e., the dealer), and suggestions by renowned great auk scholars made the birds in Bremen, Brussels, Kiel, Los Angeles, and Oldenburg the top candidates (Fuller, 1999).

In the museum industry, accurate records and archiving are obviously of high priority, with labels and registers providing vital information about the specimens (Boessenkool *et al.*, 2010; Shepherd *et al.*, 2013; Rawlence *et al.*, 2014); it therefore seems unexpected that the two bird skins could have been "lost". However, at the time, their significance as the final remnants of the species was not recognised. The story of the ending of these individuals' lives is well documented due to the efforts of English naturalist John Wolley and Cambridge University Professor Alfred Newton, who travelled to Iceland in the late 1850s and spoke directly with those who were part of the 1844 Eldey Island voyage (details from Wolley's notebook 'Garefowl books' published in Newton, 1861). What happened once the skins and their organs reached Denmark, on the other hand, is poorly recorded and remains speculative (Fuller, 1999).

In the archives of Cambridge University are the fragments of information that Newton learned of the birds. On notes dated 1861, it was recorded that Professor Reinhardt of the Royal Museum (Copenhagen) believed the skins and their organs had been purchased for the museum by Professor Eschricht of the University of Copenhagen. He is said to have taken the skins to the Congress of German Naturalists in Bremen in the autumn of 1844 (Fuller, 1999).

The connection with the skins and the Congress in Bremen could be what led Steenstrup to inform Grieve of his suspicions that the specimen at the museum in Bremen (MK134) was indeed one of the last birds (Grieve, 1885). Yet, this bird was bought by the museum at the time of the Congress from the Hamburg dealer Salmin, not Eschricht. Therefore, while the possibility may be there for Salmin to have first had the bird from Eschricht and then sold it on, it is also likely that it was a bird he had in his stocks prior to 1844 (Fuller, 1999). This study shows The Bremen Auk is not a match with either of the organs from the last pair, suggesting that it did indeed come from an earlier raid of Eldey.

The specimen in Kiel, the Schleswig–Holstein Auk (MK138), was purchased in 1844. With such a suggestive purchase date it is a contender in the mystery (Fuller, 1999). Professor Steenstrup was quoted by Grieve as saying, "*If really purchased in 1844, it might perhaps be the second of these two Garefowls got in 1844, but traditionally I never heard that mentioned*" (Grieve, 1885) (Grieve, 1885 Appendix p. 13). Our study shows this specimen was not a match, so Steenstrup was correct in his belief.

With regard to The Oldenburg Auk (MK133), this specimen was once regarded by nineteenth century scholars as belonging to one of the last birds. However, the records for this bird shows it was obtained prior to 1844 and is therefore ruled out (Fuller, 1999). It was tested in this study due to the suggestions of these early researchers but was not a match.

The history of The Brussels Auk (MK135) and Dawson Rowley's Los Angeles Auk (MK136) can be traced back to 1845 when they were said to be in the hands of a well-known, and well connected, great auk dealer, Israel of Copenhagen. Israel is known to have had excellent links with Iceland and spent his winters in Copenhagen and his summers in Amsterdam (Fuller, 1999). Fuller suggests that perhaps Israel, if he did not receive them direct from Iceland, purchased the birds in Bremen from Eschricht. The birds have a detailed history, passing through the hands of several dealers. From Israel, they were bought by Lintz, a Hamburg

merchant, and in 1845 were sold on to the Amsterdam branch of the dealer, Frank. In Newton's notes at Cambridge it was recorded that Frank believed the two skins he bought were from the last pair. The Brussels Auk was purchased in 1847 by Viscount Bernard Du Bus Ghisignies, director of the Brussels Museum (Fuller, 1999). The history of The Brussels Auk therefore strongly supports our positive match with MK131.

If the bird in Brussels, which came from Israel of Copenhagen, is from one of the last birds, then this would suggest that the second bird he had would also be from Eldey in 1844 and therefore be a positive match with the second set of organs. Israel's second bird has an even longer story than that of MK135, but it now resides in the Natural History Museum of Los Angeles County (Fuller, 1999). This specimen, Dawson Rowley's Los Angeles Auk (MK136), was tested, and the results showed it did not match LastGA2_Heart. With this negative result, we can only speculate which of the remaining untested birds could be identified as the second individual.

A possible scenario to explain the mismatch between Dawson Rowley's Los Angeles Auk (MK136) and the internal organs from the Natural History Museum of Denmark involves a mix up of skins. Dawson Rowley's Los Angeles Auk, was once one of two great auks owned by George Dawson Rowley. During the 1930s, they were passed to Captain Vivian Hewitt who owned two additional specimens. The four specimens are currently held in Cardiff, Birmingham, Los Angeles, and Cincinnati. At Hewitt's death, his collection had been put under the control of Spink and Son Ltd., a London dealer, who offered them for sale. While organising Hewitt's affairs, the four birds were mixed up. The identity of the birds now in Birmingham and Cardiff could be easily resolved, but those now in Los Angeles and Cincinnati are harder to determine. It is thought that their identities could be determined from annotated photographs taken in 1871 by George Dawson Rowley when they were in his possession (Fuller, 1999). However, we speculate that their identities were not correctly resolved and that perhaps the bird in Cincinnati was the original bird from Israel of Copenhagen. If this were the case, then it would explain why the Los Angeles bird fails to match with either of the last great auk organs held in Copenhagen.

In summary, we suggest that The Brussels Auk is the skin from the last male great auk killed on Eldey Island in June 1844. The skin of the female killed at the same time remains unaccounted for, but a common history with The Brussels Auk makes the skin currently held

at Cincinnati Museum of Natural History and Science, a likely candidate. A re-evaluation of the historical records may reveal further candidate skins amongst those currently held in museums around the world.

5.5 Conclusions

Ancient DNA has been used to evaluate museum collections in the past, albeit usually for taxonomic identification of unidentified or misidentified accessions. Our study shows an alternative use of the technology. It demonstrates the utility of molecular tools and advanced sequencing to contribute to questions, which are not primarily biological or molecular but rather historical in nature. The unravelling of the mystery surrounding the whereabouts of the skins of the last two great auks represents a fascinating element in the story of extinction and human involvement in that process.

5.6 Additional information

Nota bene: Since the publication of this paper, we have been able to secure a collaboration with the Cincinnati Museum of Natural History and Science and consequently acquired permission to sample the specimen in their collection which we speculate belongs to the female of the last pair. Thus, we hope that we will soon have a result which will either confirm the present speculation and find the Cincinnati great auk specimen to be a match, or we shall identify the possibility that a previously unrecognised specimen is a possible candidate.

5.6.1 Publication information

Thomas, J. E., Carvalho, G. R., Haile, J., Martin, M. D., Castruita, J. A. S., Niemann, J., Sinding, M.-H. S., Sandoval-Velasco, M., Rawlence, N. J., Fuller, E., Fjeldså, J., Hofreiter, M., Stewart, J. R., Gilbert, M. T. P. and Knapp, M. (2017) 'An 'Aukward' Tale: A Genetic Approach to Discover the Whereabouts of the Last Great Auks', *Genes*. Multidisciplinary Digital Publishing Institute, 8(6), p. 164. doi: 10.3390/genes8060164.

5.6.2 Author contributions

J.E.T. conceived the study, J.E.T., J.H., G.R.C., M.T.P.G., and M.K. designed the experiments; J.E.T., J.H., M-H.S.S., and M.S.-V. conducted the experiments; J.E.T.,

M.D.M., J.A.S.C., and J.N. analysed the data; N.J.R., M.H., and J.R.S. provided the initial framework for the study; all authors contributed to writing the paper.

Chapter 6: General Discussion

6.1 Thesis highlights

Despite the vast amount of literature pertaining to the great auk, at the outset of the current study, and even beyond, there remain many unanswered questions regarding aspects of their evolution and extinction. The current research has, however, allowed for a number of these questions to be addressed.

Prior to this thesis, it was hypothesised that oceanographic-related size differences existed in the great auk, with those from the North-West (NW)/Low Arctic region being larger than those from the North-East(NE)/Boreal region (Burness and Montevecchi, 1992). Our results contradict the previously published morphological evidence supporting this hypothesis, with an overall trend of no statistically significant size variation of the humerus found between great auks from the NW and those from the NE. The purpose of this analysis was to identify levels of population structure in the morphometrics that may be an indication of genetic structure and limited gene flow. Identifying such features is important as it would have had implications for their extinction risk. As we found no evidence of morphometric differences we concluded that studying the genetics of the great auk would be the most appropriate and informative course of action. Using molecular data, we would be able to determine if population structure within the great auk existed and what this may mean for their extinction.

The demise of the great auk is discussed in almost every piece of literature written about it. While the most probable cause of the great auk extinction was assumed to be due to the intense hunting of the bird that occurred from around 1500CE, it was not known if they were already in decline from other factors. It had been proposed that perhaps past climatic events had impacted the great auk, which could have contributed to their extinction (Bengtson, 1984). One of the greatest strengths of ancient DNA (aDNA) research is its ability to allow us to travel back in time, giving us an insight into the lives of species lost to us today, by examining genetic characteristics. Analysis of aDNA has been applied to several charismatic megafaunal species, in order to investigate questions such as the drivers of extinction, and to reconstruct past population demographics. The current study presents the first time that anyone has empirically investigated the cause of extinction in the great auk. By sampling material from across the great auks' former distribution, and across a time frame of 15,000ybp-170ybp, we generated a data set of 41 mitochondrial genomes, revealing several insights. Firstly, we found no evidence of population structure throughout their range, or through time. The result was somewhat

unexpected in light of the previously published morphometric results. Demographic reconstructions showed a lack of evidence for decline over the past 250,000 years, with a female effective population size of ~300,000. We thus concluded that the great auk was not especially vulnerable or at risk of extinction prior to the intense hunting that began approximately 500 years ago, and that any past climate-driven environmental change was unlikely to have caused any detectable loss of genetic diversity or bottlenecks. Population viability analysis estimated the level at which hunting would have had to occur to cause extinction in such a short space of time. It revealed that hunting of just 5-6% of the total population size would have been able to cause extinction within the time frame that it is documented to have occurred. While it has long been known that great auks were heavily hunted, it has been previously questioned how it would have been possible to cause such an abundant and wide spread species to go extinct so quickly. Consequently, it was suggested that perhaps they were in decline or possibly never particularly abundant. However, the findings of this study show not only was the great auk once an abundant species, but that there was no evidence to suggest that the great auk was in decline or would have gone extinct naturally had humans left it alone.

Following the success and interesting conclusions of the mitogenome study, we aimed to use nuclear markers to investigate the population genetics of the great auk. It is now generally accepted that using a combined approach of both nuclear DNA (nuDNA) and mitochondrial DNA (mtDNA) provides a more detailed picture of demographic changes through time. Unfortunately, due to low coverage of targeted markers, in all but one of the samples, we were unable to perform meaningful population genetic analysis. Our result was not unexpected, as for much of the history of the field of ancient DNA, studies focussed on sequencing mtDNA, as obtaining sequences of nuDNA is more challenging (for reasons discussed in Chapter 4). Nevertheless, in recent years, technological advancements, and reduced costs have led to greater success of sequencing nuDNA data and it is not unlikely that further technological improvements will allow researchers to routinely use nuclear data in ancient population genetic analyses in the not-too-distant future.

The final data chapter of this thesis, details the mystery of the missing skins from the last documented pair of great auks, killed on Eldey Island, 1844. Whilst collecting samples to address the main aim of the project, we were able to sample material which allowed us to attempt to resolve the mystery. The organs of the last pair have remained at the Natural History

Museum of Denmark since their acquisition. The skins, however, were sold and their whereabouts became uncertain. Several scholars had proposed the locations of the two skins, with one of the most comprehensive lists composed by Fuller (1999). By sampling the organs of the last birds, and samples from five candidate specimens, we set about to match organ to skins using the sequenced mitogenomes. Fortunately, the high levels of genetic diversity that can be seen in our mitogenome dataset allowed us to secure a robust match between one set of organs, those from the male, and the skin now in the RBINS in Brussels. Unfortunately, no skin was found to be a match for the female individual, however, we hope that very soon the mystery will be fully resolved.

Having presented highlights, I consider below collectively key aspects of our findings in relation to the wider literature.

6.2 The extinction of the great auk in comparison to another avian extinction

There are several avian species whose names are synonymous with the term extinction. The great auk is of course one such case. Other species include perhaps the best-known example of any extinction, the dodo (*Raphus cucullatus*), the passenger pigeon (*Ectopistes migratorius*) and the moa (Aves: Dinornithiformes). The extinction of all these species has been attributed to humans (see below for further detailed discussion). In light of our findings we now compare what has been found regarding the cause of extinction for two of these, the passenger pigeon and moa, and what the results may mean for the dodo, whose extinction is yet to be similarly explored.

The findings of the current study suggest that the great auk was not in decline prior to the intense human hunting that began in ~1500CE. Therefore, observations suggest the cause of extinction to be due to human hunting alone. For many of the megafaunal mammals whose extinction has been examined, we find that overall there is at least some contribution of environmental factors to their extinction (e.g. steppe bison (Shapiro *et al.*, 2004; Lorenzen *et al.*, 2011), cave bear (Stiller *et al.*, 2010), wild horse (Lorenzen *et al.*, 2011), Patagonian megafauna (Metcalf *et al.*, 2016)). However, as commented on in Chapter 3, the extinction of the great auk is similar to another avian extinction. The moa were large, flightless, herbivores, found throughout New Zealand (Rawlence *et al.*, 2012). Like the great auk, the extinction of

the nine species of moa was thought to have been caused by humans, but it was not known if they were already in decline before this time due to factors such as climate-driven environmental change. It has now been found that the moa showed genetic stability through time, evidence that they were not in decline and that they were able to track changes in habitat caused by environmental change (Rawlence *et al.*, 2012; Allentoft *et al.*, 2014). Just as in the case of the great auk, the extinction event occurred too quickly (for the moa in 1-2 centuries (Perry *et al.*, 2014)) to be recorded in their genetics (Allentoft *et al.*, 2014). While the extinction of the moa has been attributed to hunting and habitat destruction (Rawlence *et al.*, 2012; Holdaway *et al.*, 2014), whereas for the great auk hunting alone appears to have caused its demise, this comparison highlights two cases where humans have caused rapid extinction to species that were not previously in decline.

The passenger pigeon of North America, was once the most abundant bird species in the world (Halliday, 1980; Hung *et al.*, 2014). Its extinction is one of the best described, as the species declined from an estimated population size of 3-5 billion in the 1800s, to zero in 1914, when the last captive passenger pigeon died in Cincinnati Zoo (Johnson *et al.*, 2010; Hung *et al.*, 2014). Its extinction, while said to be caused by human exploitation, raised questions as to how an abundant bird could be driven to extinction in a relatively short period, similar to the questions tackled herein. The passenger pigeon had a long history of being hunted for food, by both Native Americans, and European colonists (again similar to the great auk) (Halliday, 1980). In the mid-19th century, hunting also came in the form of commercial hunting for meat and sport shooting, and this coincided with human impacts to their habitat (Stanton, 2014). In 2014, Hung et al. used pairwise sequentially Markovian coalescent (PSMC) to estimate the effective population size changes through time, using the genomes of three passenger pigeons. They reported that the passenger pigeon had not always been as abundant as it was in the 1800s, but experienced drastic fluctuations in population numbers, resembling those of an 'outbreak' species (Hung *et al.*, 2014). The combined approach of investigating the demographic history using aDNA, coupled with environmental niche modelling, led Hung et al. (2014) to conclude that the period of human exploitation coincided with a naturally occurring reduction in population size, caused by changes in ecological conditions (year to year variations in their food source carrying capacity), therefore causing their rapid demise (Hung *et al.*, 2014). However, in a recent study by Murray et al. (2017), in which a Bayesian skyline model was used to infer the passenger pigeons' population dynamics (using 41 mitogenomes), it was reported that the population size was high and stable for a long period of time prior to the period

of intense hunting. This finding by Murray et al. (2017) contradicts that of Hung et al. (2014). Murray et al. (2017) also investigated the impact of natural selection on the passenger pigeons' genomes through comparative analysis with its closest living relative, the band-tailed pigeon. They suggest that the passenger pigeons' large population size allowed for faster adaptive evolution, thereby allowing for the removal of harmful mutations. The results suggested that the passenger pigeon may have evolved traits that were adaptive when their population was large, but that made it more difficult for them to survive after the population numbers declined due to hunting (Murray *et al.*, 2017). This new result provides a more similar picture of extinction to the great auk, and even that of the moa, than that suggested by Hung et al. (2014). All three species (great auk, passenger pigeon and moa) appear to have had large, stable population numbers for a significant period of time and do not appear to have been in decline or at risk of extinction prior to the respective periods of intense hunting by humans. All three species extinctions have therefore been attributed to humans and provide stark warnings for the exploitation of extant species.

An additional contributing factor that has been suggested for the passenger pigeon, may also be relevant when discussing the extinction of the great auk. The passenger pigeon, was a very social bird, feeding and breeding in large groups as discussed above. The great auk was also a species that exhibited highly social behaviour, as evidenced from the large numbers that bred on Funk Island, and reports of rafts of seabirds observed in wintering months (Fuller, 1999). There are many advantages of living in large colonies, for example, reducing the effect of predation, and increased chances of finding a mate (Schippers *et al.*, 2011). Additionally, it appears that for the passenger pigeon at least, a large population size allowed for faster adaptive evolution and removal of harmful mutations (Murray *et al.*, 2017). Halliday (1980) suggested that if the colony size of the passenger pigeon fell below a critical point, they would be unable to form sufficiently large breeding colonies, and this would impede reproductive rate/success and the species would become extinct (Halliday, 1980). Halliday (1980) noted that it would be possible for colony size to fall below the critical value while the species still appeared to be quite common. With regards to the great auk, Fuller (1999) (page 63) stated *'Once its* [the great auks] *numbers had fallen below a certain level, no matter that the number was still very large, the species was not viable and was doomed'*. Montevecchi et al. (2007) likewise state that hunting in the 1700s caused a decline in population numbers to a point below the minimum of a viable level. Such trends describe a phenomenon known as the Allee Effect, which is a defined as a casual positive relationship between individual fitness and either population size

or density (Stephens *et al.*, 1999; Courchamp *et al.*, 2006). Therefore, in addition to the impact that hunting was having in simply reducing the population levels and affecting reproductive success, an Allee Effect would have meant that numbers might not have needed to become particularly low to ensure extinction, of the great auk or passenger pigeon, was inevitable.

Like the great auk, much has been written about the dodo, yet there remain many unanswered questions with regards to its evolution, life history and extinction (Hume, 2006). The dodo, was a large, flightless bird, endemic to the island of Mauritius (Halliday, 1978). Ancient DNA has been used to confirm its taxonomy, placing it in the Columbidae family, as a sister taxon to the extinct solitaire, and its closest living relative being the Nicobar pigeon (Shapiro *et al.*, 2002). The extinction of the dodo has yet to be investigated empirically using methods such as those employed here, thus the cause is much less well understood. It is possible that investigating the extinction of the dodo may prove to be more challenging that the great auk, moa or passenger pigeon, as dodo samples are from hot, humid environments. Such conditions can lead to poorly preserved DNA (as discussed elsewhere in this thesis), however, aDNA has been extracted from dodo remains (Shapiro *et al.*, 2002) and an increasing number of studies are obtaining DNA sequences from samples of tropical climates. Therefore, it may not be long before such a study is conducted. As with the other examples, the extinction is likely to be due to anthropogenic causes. However, hunting alone is less likely to be the cause (although evidence of hunting does exist (Janoo, 2005)), but more due to introduction of commensal animals (such as rats and pigs) which increased competition for food and direct predation on chicks and eggs, and habitat destruction (Hume, 2006). Several accounts detail the decline in population numbers and its increasing rarity (Hume, 2006). The dodo did possess several traits thought to make a species more susceptible to extinction (McKinney, 1997). It was, large, flightless, had restrictive habitat preferences, very small range/island endemic, and showed an intolerance to humans. It is well documented that island endemics have higher extinction rates than nonendemic species, possibly due to the chance of increased levels of inbreeding depression (Frankham, 1998). Therefore, irrespective of human arrival to Mauritius, was the dodo already heading for extinction? The examples discussed above, show that in at least two (great auk and moa), if not all, then the species could still be in existence if humans had not hunted them in the way that they did. Investigating the extinction of the dodo, using aDNA and demographic reconstructions may show that like the moa or great auk, populations were stable and they were not in decline. As an island endemic, one which potentially faced impacts

of extreme environmental conditions, such as cyclones which caused food shortages (Hume, 2006), which may lead to population declines and potentially bottleneck/expansion cycles, it is likely that genetic diversity was never particularly high and that they were already in decline. In such a case, the arrival of humans, and subsequent hunting/invasive species etc., may have accelerated their extinction but not have been the sole cause.

6.3 Could we have predicted the extinction of the great auk?

Preventing species decline and extinction is one of the main driving forces behind conservation practices. Thus, monitoring species and assessing them for their risk of extinction is common practice (e.g. the IUCN Red List). When assessing a species threat level factors evaluated include population size, decline rate, range size, life history/ecological/behavioural traits and threats such as changes in habitat, catastrophes etc. (Mace *et al.*, 2008). It has been said that should genetic factors be ignored, then the risk of extinction will be underestimated. Furthermore, if said genetic factors are ignored, then the extinction risk cannot be considered scientifically credible (Frankham, 2005). Such genetic factors which could be investigated and thus used in the assessment of a species risk to extinction include, levels of inbreeding, genetic diversity, gene flow, genetic effective population size and population structure (Frankham, 2005; Frankham *et al.*, 2014). It has also been said that most species are not driven to extinction before genetic factors have time to impact (Frankham, 2005).

The great auk has been extinct for ~170 years, yet if it was still extant, would we have been able to predict, and even prevent, its extinction? If we were to apply the type of assessment described above, we could make the following conclusions from the results of this study and previously published information. Great auks' population size was large, possibly in the millions, therefore it would be described as an abundant species. It was widely distributed around the North Atlantic, breeding on isolated islands, though with apparent gene flow across its range. While human populations have spread, many of the islands that the great auks once bred on still remain unpopulated, and are in fact breeding bird sanctuaries, with very restricted access (such as Eldey Island and Funk Island). Threats to the great auk would be the same as those faced by its extant relatives', and as it shared many of the same life history traits (with the exception of flightlessness) the risk from the threats would likely be the same. Additionally, and perhaps most importantly, if the great auk was extant today, and if we had investigated its population genetics, then we would have found no evidence from their genetics to suggest they

were in decline, inbreed, reproductively isolated etc. i.e. at risk of extinction. Thus, if the great auk was alive today, then while it would have been monitored, it is unlikely that it would be classed as threatened/endangered.

The results of our study and this hypothetical assessment have led us to consider implications for extant species. Our results show it is possible to cause extinction to an abundant, wide spread, genetically healthy species, in a relatively short period of time. Therefore, while conservation practices exist today to monitor species risk, if a species is not regarded as threatened due to having the characteristics described above, then its extinction risk may be underestimated and they may not be monitored as closely as a species known to be in decline or at risk of extinction. Such a result may be especially relevant for species which are data deficient or perhaps yet to be discovered.

A similar hypothetical assessment of the extinction of the passenger pigeon has been published. Stanton (2014), used a series of population models, which took in to account various contributions to the decline of the passenger pigeon, in addition to life history factors, to assess if present-day risk assessment would have predicted its extinction. Findings suggested that if the species had been listed as threatened, based on rates of decline, then it could have been detected sufficiently early for conservation practices to present extinction (Stanton, 2014). Stanton (2014), also noted that with regards to simulations which included commercial harvest, only a small proportion of replicates caused the species to become critically endangered or extinct. She therefore suggested that it was not the commercial hunting that was unsustainable, *per se*, but rather the intensity and the manner in which it was carried out that was the problem (Stanton, 2014) (i.e. hunting breeding/nesting birds (Halliday, 1980)). Such assertions are perhaps similar to our simulations of great auk hunting in which we show that harvesting the species at certain levels would not necessarily cause extinction or even a decline. Therefore, if the numbers of great auk killed annually had been monitored, as in hunting of alcids today, and protection laws (or at least effective laws) put in place, then the commercial harvest could have existed but the extinction may not have occurred. Stanton (2014) suggested that the commercial harvest of passenger pigeons was at a level at which there was extremely high levels of waste, with birds left uncollected and decomposing on the forest floor. We also see this with great auk hunting, with more birds killed that could be collected and the manner in which the hunting occurred being destructive and wasteful (Fuller, 1999; Gaskell, 2000). Therefore, as Stanton (2014) suggests with the passenger pigeon, there would have been room

to increase the efficiency of the harvest, i.e. reduce the numbers killed to levels which did not leave waste, but with minimal impact to the financial gain (Stanton, 2014).

6.4 Implications of findings for extant seabirds

The great auk is of course not the only seabird to have been impacted by humans, prehistorically, historically and even today. Indeed, seabirds are more threatened than any comparable groups of birds, with a third of seabird species at risk of extinction, and one half are in or likely to be in decline (Croxall *et al.*, 2012). Threats to seabirds have increased in line with human population growth. Major threats include, climate change, habitat loss, introduced species, energy production, pollution, overfishing of prey species and fishing industry related impacts such as bycatch and entanglement and also direct exploitation in the form of harvesting eggs, chicks and adult birds (Croxall *et al.*, 2012; Paleczny *et al.*, 2015). Since seabirds play a globally important role in ecosystems, both as indicators for marine ecosystem health and function, localised or global declines are of interest in conservation, as well as wider aspects of the marine environment (Paleczny *et al.*, 2015). While contemporary seabird harvesting is monitored and with some protection laws in place, synergistic impacts of global warming, overfishing of prey, pollution from oil spills and habitat loss from our growing populations, mean that ongoing conservation strategies remain a high priority.

As discussed, species monitoring and assessing risk of extinction is common place today. However, we have shown, that it is possible to cause extinction of an abundant, wide spread and genetically healthy species, in around 300 years if the hunting level is high enough. With many of the extant relatives of the great auk listed on the IUCN Red List of threatened species, we shall discuss what our results mean for two of them.

The IUCN considers the razorbill (closest living relative of the great auk (Moum *et al.*, 2002)) to be 'Near Threatened'. European population numbers are estimated to be ~1million adult birds, which is thought to be around 95% of the global population (BirdLife International, 2016a). Threats to the species includes the impacts of climate change such as changes in temperatures, and shifts and reductions to prey, and the unregulated hunting in a number of places including Labrador, Newfoundland and Greenland (BirdLife International, 2016a). Population genetic analyses of the razorbill have found genetic diversity and genotype

frequencies, suggestive of restrictions to gene flow (Moum & Árnason, 2001). Moum & Árnason (2001), suggested that their results indicated that razorbills originated from refugial population in the south-western Atlantic Ocean, through sequential founder events and expansion to the east and north (Moum & Árnason, 2001). Therefore, the razorbill has shown several indications of an increased risk of extinction compared to the great auk. Hunting simulations performed in our study showed that for a population of 1 million birds, harvesting at levels greater than 5% could cause extinction. Such a result may be very relevant and important for the razorbill. In Iceland, an average of ~24,000 razorbills were killed annually between 1995-1999 (Petersen, 2004). While this may be only 2% of the current estimated adult population, this is only in one part of the razorbills range. Therefore, when combined with hunting in other areas, some of which is unregulated and illegal, levels may be approaching the tipping point of having detrimental impact on population numbers. As the razorbill is showing a decline in numbers, poor genetic health and evidence that it is impacted by climate change, the numbers needed to be harvested may be a lot lower than for the great auk. With unregulated hunting in much of its range this is therefore concerning and should be addressed.

The thick-billed murre has been classified as 'Least Concern' (BirdLife International, 2016b). Murres are one of the most abundant seabirds in the northern hemisphere, with population numbers exceeding 20 million (BirdLife International, 2016b). While the IUCN reports numbers are increasing, other monitoring systems suggest a downward trend across parts of their range over the last thirty years (Ganter & Gaston, 2013). The IUCN Red List lists the major threats to thick-billed murres as incidental kills form fisheries, competition of commercial fisheries, oil pollution and climate change (BirdLife International, 2016b). Population genetic studies of the thick-billed murre have found various levels of population structure, with the identification of four genetically differentiated groups, but little population structure within the Atlantic (Tigano *et al.*, 2015). With regards to hunting off the coast of Newfoundland and Canada, between 200,000-300,000 murres are killed legally every year. This number was even higher before the mid-1990s, when between 300,000-700,000 thick-billed murres were being harvested annually (Wilhelm *et al.*, 2008). If population numbers are in excess of 20 million, then clearly huge numbers (1-1.2million) would need to be harvested, if as with the great auk a harvest of 5-6% could lead to extinction. However, we know that it is possible, from our great results, that species' do not necessarily have to show the traits indicative of a vulnerable species, to go extinct when the numbers are not monitored. If the thick-billed murre continued to decline then perhaps it may reach a level where harvesting 5%

of the population would not be such a high number. It may be possible for such a decline to occur without even being noticed or in a short period of time, if for example, an oil spill killed a huge proportion of the population. The overall inference from such comparisons is that species today are in a more fortuitous position than the great auk in that they are monitored and protection laws exist. The key message is that even a species which does not appear to be at risk due to high population numbers, wide distribution etc. can become extinct rapidly, if effective/ regular monitoring and conservation practices are not enforced.

6.5 Future research

Our study has reiterated the utility of aDNA in investigating the lives of extinct species. While it has revealed insights into the life and extinction of the great auk, and even events that have occurred following their demise, it has also identified other priority questions and areas for future work.

Future work already hinted at, aside from expanding the mitogenome study to include more individuals, is to obtain sequences from the nuclear DNA. As discussed, the study may be improved by making methodological modifications or employing more recently published methods, which could increase our changes of successfully sequencing great auk nuDNA. Additionally, while the main scope of this project has been to use genetics, it would be informative to expand the morphometric study to include other bones of great auks, especially those not from the wing, to determine if the pattern we observed persists.

An interesting and fairly recent development in the field of aDNA, due to technological advances, is the study of microbiota and microbiomes from ancient samples- palaeomicrobiology (Warinner *et al.*, 2014). Studies have used paleofaeces (Tito *et al.*, 2008), coprolites (Tito *et al.*, 2012), dental calculus (Adler *et al.*, 2013; Warinner *et al.*, 2014), and more recently, soft tissue of frozen mummified humans, (Lugli *et al.*, 2017), to reveal insights into the microbiomes of humans and used this information to better understand processes of our evolutionary past. With regards to the great auk, we have an ideal opportunity to study the gut microbiome of an extinct seabird because of availability of organs stored at the Natural History Museum of Denmark. The organs have the potential to provide an insight into the diet, diseases, and microbial communities of the great auks, which can be compared with living relatives. Learning more about the microbiome of a species also has important implications

relating to a point raised in the 'General Introduction' (Chapter 1), de-extinction. A species' microbiome plays an important role in the species survival. Therefore, it is something which cannot be overlooked when discussing a species potential as a candidate of de-extinction. One significant challenge, aside from the initial technical issues and ethical considerations, is how one would recreate an ancient microbiome and subsequently equip the de-extinct species with a fully functional microbiome (Richmond *et al.*, 2016). While sequencing the microbiome of the great auk does not mean that we would ultimately be able to recreate it, it is the first step in a potentially vital process should we one day have the ability to bring the great auk back to life.

While the current study has examined the genetics of the species as a whole, it may also be of interest to investigate the relatedness of samples to determine the relationship between the bones, mounted specimens in collections, and even between the specimens and eggs. Due to the high genetic diversity that allowed us to determine a match in the 'Skin Mystery', it could be possible to use mtDNA to look at maternal relationships. However, if we improve the methods for obtaining nuDNA (as discussed in Chapter 4), perhaps a study similar to Allentoft et al. (2015) which used microsatellite analysis to identify relatives between moa bones, may be possible. As we found no evidence of population structure it is impossible to determine a specimens' origin to geographic location. However, as some mounted specimens do have provenance, it would be worthwhile exploring the relationship between individuals collected from the same or different locations, in terms of relatedness, as such findings may provide information for community structure, breeding systems etc.

Similarly, it may also be of interest to determine the sex of bones (as in (Allentoft *et al.*, 2010)), and mounted specimens. The great auk is not thought to be sexually dimorphic, however, it has been suggested that sex could be identified through plumage, with females apparently possessing a lilac/grey fringe of feathers on their flanks (proposed by Rothschild (1907) discussed in (Fuller, 1999)). More recently, we know of a researcher attempting to determine sex of the great auk by using beak morphology. Therefore, such genetic information would be complimentary to such projects or again reveal valuable and new information for their biology and social behaviour.

In addition to the great auk demographic reconstruction that has been performed in this study, it would also be interesting and informative to determine if the movements that we see were related to changes in prey. Two possible ways to do this could be to (i) use stable isotopes, or

(ii) ecological niche modelling (similar to Hung et al. 2014 for the passenger pigeon). The use of stable isotopes ($\delta^{13}C$ and $\delta^{15}N$ ratios) would provide further insight into species ecology (i.e. diet as previously investigated (Hobson & Montevecchi, 1991)) but it may also show changes that accompanied climatic variations or differences in climatic or temporal range. This may provide further insight into various aspects of the great auks' biology, population dynamics or extinction. For example, we may be able to detect dietary changes, that are associated with movements caused by climatic changes.

Clearly, there is still much to be learned about the great auks' evolution and extinction, however, the research presented here has provided an interesting insight into their lives, and the suggestions of future work have the potential to reveal a great deal more.

Appendices

Appendix 1-Protocols

Morphometric sampling protocol

1. Samples inspected for damage and archaeological important features such as cut marks.
2. Samples were chosen to ensure only individuals were sampled, for example- only right humerus or only one bone from each layer of each site etc.
3. Sample information recorded: institution number, age, site information, relevant comments, element (type of bone).
4. Photograph taken prior to measuring/sampling for DNA on 1 cm squared graph paper.
5. Measurement data collected using digital callipers according to Hufthammer (1982) (Chapter 2: Supplementary Materials) and recorded.

Protocol for collecting bone material for DNA extraction

Following the collection of morphometric data, bones were sampled to provide material to be used in DNA extractions using the following protocol.

Notes: Sampling was conducted in an area deemed to be most sterile but this varied between institutions. Sampling was usually performed in back room of museum/ collection stores, but never where amplification of modern DNA had occurred. Similarly, sampling was only carried out on days when there had been no contact with amplified DNA. Prior to sampling the work area was made as sterile as possible and gloves and face mask were worn while handling bone and sampling. Tubes were labelled with institute number, lab ID and date prior to sample collection.

1. Foil placed over work surface and weighing boats wrapped in foil.
2. Bones examined and site of sampling decided (sites were chosen so to avoid any morphologically informative features and historical marks, such as cut marks)
3. Initially, the surface of the bone was lightly drilled to remove the out layers of bones that were most likely to be contained with environmental DNA. This material was discarded.
4. A Dremel drill fitted with a Dremel 107 2.4 mm engraving cutter or Dremel cutting wheel was used to collect sample material, depending on institutes preference.

5. The ‘cleaned’ sample site was then drilled to collect material, which was caught in the weighing boat.
6. Amount of material collected varied, depending on institute preference, but where possible at least 50-100 mg of bone was collected for DNA extraction.
7. Bone powder/bone section was transferred from the weighing boat to labelled tube and stored initially at room temperature but once in lab at -20/-80ºC.
8. Between each sample, drill was cleaned to remove bone dust, foil was replaced on surface, gloves changed, new weighing boat used and covered in clean foil and drill bit was changed.

Protocol for collecting tissue material for DNA extraction

Samples to be used in the skin mystery were collected from the candidate list proposed by Fuller (1999) and the organs of the two individuals killed on Eldey Island, 1844. Sampling was performed using sterile equipment and the appropriate method for sample type, which cause minimal physical damage to the specimen. Gloves were worn whilst handling the specimens and for sampling.

With the exception of samples MK131, MK132 and LastGA2_Heart, sampling of mounted specimens to collect feather or tissue samples was performed by the respective museum curators. Instructions were sent to curators on how to perform sampling. If tissue from the body or toepad was to be collected a sterile scalpel was used to cut a 1x1cm square (if possible) of tissue, which was placed into a small Ziploc bag. If feathers were collected then these were plucked from the body of the bird, from the layers closest to the body and if possible, retained body tissue on quill.

To sample the heart and oesophagus, organs were removed from the respective jars (one at a time) and a section of tissue was cut from a site to cause as little damage as possible, using a sterile scalpel. For the oesophagus, tissue was removed from the end and for the heart, tissue removed from the end of the aorta.

DNA extraction from bone
Dabney et al. (2013)

Buffers

Extraction buffer [10mL] ***(irradiate with UV before usage)*** **Make fresh every time**

Reagent	Volume	Final concentration
Water	745 µl	
0.5 M EDTA, pH 8.0	9 ml	0.45 M
10 mg/ml Proteinase K	250 µl	0.25 mg/ml
Tween 20	5 µl	0.05%

TET buffer [50mL] ***(irradiate with UV before usage)***

Reagent	Volume	Final concentration
Water	~49.4 ml	
0.5 M EDTA, pH 8.0	100 µl	1 mM
1 M Tris-HCl, pH 8.0	500 µl	10 mM
Tween 20	25 µl	0.05%

Binding buffer [BB] [50mL] ***(irradiate with UV before usage)*** **lasts 1 month**

Reagent	Volume/amount	Final concentration
Guanidine hydrochloride	23.88g	5 M
Water	to 30 ml	
Isopropanol	to 50 ml	40%
Tween 20	25 µl	0.05%

Add salt to a falcon tube. Fill up with water to 30 ml (using the graduation of the falcon tube is sufficient). Mix to dissolve the salt (if necessary heat it in the microwave very briefly). Fill up to 50 ml with isopropanol. Add Tween 20 last.

Day 1

Make Extraction Buffer

Extraction

1. In a 2.0 ml SafeLock LoBind tube, add 1 ml of extraction buffer to up to 50 mg of sample powder.
2. Suspend the sample powder by vortexing if necessary and rotate the tube overnight at 37°C in the incubator.

Make sure you have enough Binding Buffer for day 2.

Day 2

1. For each sample (and control), transfer 10 ml Binding Buffer to a 50 ml falcon tube and add 400µl 3M sodium acetate (or make a premix, e.g. by adding 2 ml acetate to 50 ml Binding Buffer, and take ~ 10.4 ml of the premix). Using the graduation on the wall of the falcon tube is sufficiently precise.
2. Spin the tubes containing sample and extraction buffer for 2 min at maximum speed in a table centrifuge to pellet residual sample powder.
3. Transfer the supernatant to the falcon tube containing the Binding Buffer / sodium acetate mix. Mix gently by shaking. Save the bone pellet for later experiments (digest).
4. Label the cap of a MinElute spin column and slightly distort the inner rim of the MinElute cap using forceps.
5. Force an extension reservoir of a V-spin column into the opening of the MinElute tube. Remove the extension reservoir/MinElute assembly from the collection tube and place it into a 50ml falcon tube.
6. Pour the sample/Binding buffer mixture from step 5 into the extension reservoir and close the falcon tube with a screw cap. Centrifuge for 4 min at 1,500 rpm in a plate centrifuge. Turn the tubes by 90°C and centrifuge for 2 more min at 1,500 rpm.
7. Remove the screw cap and place the extension reservoir/spin column assembly back into the collection tube. Carefully remove the extension reservoir and close the cap of the spin column. Close the falcon tube with a screw cap and safe the flow through for later experiments. Proceed to step 10.
8. Perform a dry spin for 1 min at 6,000 rpm in a table-top centrifuge. Discard the flow-through.
9. Add 750 µl PE buffer, spin at 6,000 rpm for 30 sec and discard the flow-through. Repeat this step.
10. Turn the spin column by 180°C and perform a dry spin for 1 min at maximum speed (13,200 rpm). Transfer the spin column into a fresh collection tube (or a 1.5 ml tube with the cap ripped off).
11. Add 25 µl TE on top of the silica membrane and let the tube stand for 2-5 min. Spin for 30 seconds at maximum speed. Repeat this step.
12. Store the elution's in freezer at -20°C until required.

DNA extraction from feather and tissue samples

Gilbert et al. (2007)

Day 1

1. Prepare samples (Feathers: cut up quill using a sterile scalpel. Tissue: cut up using a sterile scalpel.) Place in Eppendorf.
2. Make digestion buffer.

Digestion buffer								
Reagent	**Initial conc.**		**Final conc.**		**For 1 sample**		**Mix x 10 samples**	
	#	unit	#	unit	#	unit	#	unit
Tris-HCl (pH 8.0)	1	M	10	mM	10	µl	100	µl
NaCl	5	M	10	mM	2	µl	20	µl
SDS	10	% w/v	2	% w/v	200	µl	2000	µl
$CaCl_2$	1	M	5	mM	5	µl	50	µl
EDTA (pH 8.0)	0.5	M	2.5	mM	5	µl	50	µl
DTT	1	M	40	mM	40	µl	400	µl
Proteinase K	100	%	10	%	100	µl	1000	µl
H_2O					638	µl	6380	µl
Total volume					1000	µl	10,000	µl

3. Add 1000 µl digestion buffer to sample.
4. Vortex sample and incubate overnight in rotator at 55^{o}C.

Day 2: continue from Day 2 in Dabney et al. (2013) described above.

DNA extraction from heart tissue

Campos & Gilbert (2012)

Day 1

1. Cut up tissue sample using sterile blade. Place in Eppendorf.
2. Make digestion buffer.

Digestion buffer								
Reagent	**Initial conc.**		**Final conc.**		**For 1 sample**		**Mix x 3 samples**	
	#	unit	#	unit	#	unit	#	unit
Tris-HCl (pH 8.0)	1	M	10	mM	10	µl	30	µl
NaCl	5	M	10	mM	2	µl	6	µl
SDS	10	% w/v	2	% w/v	200	µl	600	µl
$CaCl_2$	1	M	5	mM	5	µl	15	µl
EDTA (pH 8.0)	0.5	M	2.5	mM	5	µl	15	µl
DTT	1	M	40	mM	40	µl	120	µl
Proteinase K	100	%	10	%	100	µl	300	µl
H_2O					638	µl	1914	µl
Total volume					1000	µl	3000	µl

3. Add 1000 µl digestion buffer to sample.
4. Vortex sample and incubate overnight in rotator at 55ºC.

Day 2

1. Centrifuge digestion mix for 3-5 min at high speed (10000xg).
2. Remove supernatant and put into new tube.
3. Add 5 x volumes of PB and mix.
4. Remove 700 µl of mixture and add to QIAquick spin column.
5. Centrifuge for 1 minutes at 6,000xg.
6. Discard waste.
7. Add the remaining sample/PB mix to the spin column and re centrifuge.
8. Discard waste.
9. Add 500 µl wash buffer PE to filter.
10. Centrifuge for 1 minute at 10,000xg.
11. Discard waste and repeat wash step.
12. Dry spin-3mins at max speed.
13. Put spin column in 1.5 ml tube. Add 62 µl EB to centre of filter and leave at room temperature for 5 minutes.
14. Centrifuge for 1 minute at max speed.

Single-stranded library build method
Gansauge & Meyer (2013), with modifications as described in Bennet et al. (2014)

Reagent setup

To make enough for 25 reactions.

Bead-binding buffer (10 ml) NB: No shelf life once SDS added.

Reagent	Volume
Water	7.63 ml
NaCl 5M	2 ml
Tris-HCl 1M (pH 8)	100 µl
EDTA 0.5M	20 µl
Tween 20	5 µl
SDS 20% (ADD IMMEDIATELY BEOFRE USE)	250 µl

Wash buffer A (50 ml) Stored safely at room temp for 1month

Reagent	Volume
Water	47.125 ml
NaCl 5M	1 ml
Tris-HCl 1M (pH 8)	500 µl
EDTA 0.5M	100 µl
Tween 20	25 µl
SDS 20%	1.25 ml

Wash buffer B (50 ml) Stored safely at room temperature for 1 year

Reagent	Volume
Water	48.375 ml
NaCl 5M	1 ml
Tris-HCl 1M (pH 8)	500 µl
EDTA 0.5M	100 µl
Tween 20	25 µl

Stringency wash buffer (50 ml) Stored safely at room temp for at least 1 month

Reagent	Volume
Water	49.5 ml
SDS 20%	25 0µl
SSC buffer 20x	250 µl

Stop solution (100 µl) Stored safely at room temp for at least 1 year

Reagent	Volume
EDTA 0.5M	98 µl
Tween 20	2 µl

TE buffer (50 ml) Stored safely at room temp for at least 1 year

Reagent	Volume
Water	49.4 ml
Tris-HCl 1M	500 µl
EDA 0.5M	100 µl

TET buffer (50 ml) Stored safely at room temp for at least 1 year

Reagent	Volume
Water	49.375 ml
Tris-HCl 1M	500 µl
EDA 0.5M	100 µl
Tween 20	25 µl

Double-stranded adapter in PCR tube, resulting solution stored at -20ºC for at least 1 year

Reagent	Volume
TE buffer	9.5 µl
NaCl 5M	0.5 µl
CL53 500µM	20 µl
CL73 500µM	20 µl

Incubate the reaction mixture in thermal cycler for 10s at 95ºC and slowly decrease the temp at the rate of 0.1ºC per second until reaching 14ºC. Add 50 µl of TE buffer to the hybridised adapter to obtain a concentration of 100 µM in total volume

Day 1

Uracil excision and DNA cleavage at abasic sites (approx. timing 1.5h)

1. Prepare the following reaction mixture for each sample and blank in 0.5ml tube. Mix by flicking and spin briefly.

Reagent	**Volume(μl)**	**Final conc. in reaction (this step/step 5)**
Water (to 42 μl)		
CircLigase buffer II (10x)	8	1.9x/1x
$MnCl_2$ (50mM)	4	4.8mM/2.5mM
DNA sample (max 29 ul)		
Endonuclease VIII (10 U μl^{-1})	0.5	0.12 U μl^{-1}

2. Incubate in thermal cycler with heated lid for 1 hour at 37^{o}C

Dephosphorylation and heat denaturation (approx. timing 20 min)

3. Add 1μl of FastAP (1U) to each reaction mixture prepared in step 1 and mix by flicking. Spin briefly. Total volume now 43 μl.

GET ICE.

4. Incubate the samples in thermal cycler for 10minutes at 37^{o}C and then 95^{o}C for 2 minutes.

While thermal cycler is still at 95^{o}C, quickly transfer the tubes into an ice water bath and let cool for at least 1 minute. Spin briefly and place back in rack at room temperature.

Ligation of the first adapter (approx. timing 4 hours)

5. Add the following to the sample mix to give final volume of 80ul. Mix by vortexing before adding circligaseII and then flick once added. Spin briefly.

Reagent	**Volume (μl)**	**Final conc. in reaction**
PEG -4000 (50%) (NB viscous)	32	20%
Adapter oligo CL78 (10uM)	1	0.125 uM
Circligase II (100 U μl^{-1})	1	

6. Incubate the samples in thermal cycler with heated lid for 3 hours at 60^{o}C.

7. Add 2 μl of stop solution to each sample. Mix by vortexing and spin briefly. (Ligation products can be stored safely at -20^{o}C for several days).

Day 2

Immobilisation of ligation products on beads (approx. timing 1 hour)

8. Resuspend the stock of MyOne C1 beads by vortexing. For each sample, transfer 20 µl of bead suspension into a 1.5 ml tube.

Add SDS to bead binding buffer (250 µl)

Heat PCR machine to 95ºC.!

Pellet the beads using magnetic rack, discard the supernatant and wash the beads **twice** with 500 µl of bead-binding buffer.

Resuspend the beads in 250 µl bead-binding buffer

GET ICE.

9. Incubate the ligation reactions from step 7 (final product from Day1) for 1 min at 95ºC in thermal cycler with heated lid.

While thermal cycler is still at 95ºC, quickly transfer the tubes into an ice water bath. Let the reaction cool down for at least 1 minute and spin briefly.

Add ligation reactions (step 7) to the bead suspensions prepared in step 8.

10. Rotate the tubes for 20 minutes at room temperature.

11. Spin tubes briefly. Pellet the beads using a magnetic rack and discard the supernatant.

Wash the beads **once** with 200 µl of wash buffer A. pellet, discard

Wash the beads **once** with 200 µl of wash buffer B.

Primer annealing and extension (approx. timing 2 hours)

12. Prepare a master mix for the required number of reactions (47 µl per reaction). Make while samples rotating

Reagent	**Volume (µl)**	**Final conc. in reaction**	**x 9**
Water	40.5		364.5
Isothermal amplification buffer (10x)	5	1x	45
dNTP mix (25mM each)	0.5	250µM each	4.5
Extension primer CL9 (100µM)	1	2µM	9

13. Pellet the beads using a magnetic rack and discard the wash buffer.

Add 47 µl of the mater mix prepared in step 12 to the pelleted beads and resuspend the beads by vortexing.

Incubate the tubes in a thermal shaker for 2 mins at 65ºC. GET ICE

Place tubes in ice-water bath for 1min and then immediately transfer tubes to thermal cycler precooled to 15ºC (leave thermal cycler lid open).

While tubes are in thermal cycler, add 3 µl of Bst 2.0 polymerase (24U) to each reaction mixture.

Mix tubes by vortexing briefly and return to thermal cycler.

14. Incubate the samples at 15ºC for 30mins.

15. Spin the tubes briefly. Pellet the beads using a magnetic rack and discard supernatant.

Wash the beads **once** with 200 µl of wash buffer A. pellet, discard.

Resuspend the beads in 100 µl of stringency wash buffer and incubate the bead suspensions for 3 min at 45ºC in a thermal shaker.

Pellet the beads using a magnetic rack and discard the supernatant.

Wash the beads **once** with 200 µl of wash buffer B.

Blunt end repair (approx. timing 1 hour)

16. Prepare a master mix for the required number of reactions (99 µl per reaction)

Reagent	**Volume (µl)**	**Final conc in reaction**	**x9**
Water	86.1		774.9
Buffer Tango (10x)	10	1x	90
Tween 20 (1%)	2.5	0.025%	22.5
dNTP (25 mM each)	0.4	100uM each	3.6

17. Pellet the beads using a magnetic rack and discard the wash buffer.

Add 99 µl of the reaction mix from step 16 to the pelleted beads and resuspend by vortexing.

Add 1 µl of T4 DNA polymerase (5U). Mix tubes by vortexing briefly.

18. Incubate the reaction mixtures for 15 mins at 25ºC in thermal shaker. Keep beads suspended during the incubation (Vortex bead suspensions every 5 min and place back in incubator immediately)

19. Add 10 µl of EDTA (0.5M) to each reaction mixture and mix by vortexing. Pellet the beads using a magnetic rack and discard the supernatant.

Wash the beads **once** with 200 µl of wash buffer A. pellet, discard.

Resuspend the beads in 100 µl of stringency wash buffer and incubate the bead suspensions for 3 min at 45ºC in a thermal shaker.

Pellet the beads using a magnetic rack and discard the supernatant.

Wash the beads **once** with 200 µl of wash buffer B.

Ligation of second adapter and library elution (approx. timing 2hours)

20. Prepare a master mix for the required number of reactions (98µl per reaction).

Reagent	**Volume** µl	**Final conc. in reaction**	**x9**
Water	73.5		661.5
T4 DNA ligase buffer 10x	10	1x	90
PEG- 4000 (50%)	10	5%	90
Tween 20 (1%)	2.5	0.025%	22.5
Double stranded adapter (100uM)	2	2.5uM	18

21. Pellet the beads using a magnetic rack and discard the wash buffer.

Add 98 µl of the reaction mix from step 20 to the pelleted beads and resuspend the beads by vortexing.

Add 2 µl of T4 DNA ligase (10U). Mix by vortexing briefly.

22. Incubate the reaction mixtures for 1 hour at room temperature. Keep the beads suspended during incubation (as described in step 18).

23. Pellet the beads using a magnetic rack and discard the supernatant.

Wash the beads **once** with 200 µl of wash buffer A. pellet, discard.

Resuspend the beads in 100 µl of stringency wash buffer and incubate the bead suspensions for 3 min at 45ºC in a thermal shaker.

Pellet the beads using a magnetic rack and discard the supernatant.

Wash the beads **once** with 200 µl of wash buffer B.

24. Pellet the beads using a magnetic rack and discard the supernatant.

Add 25 µl of TET buffer to the pelleted beads, resuspend the beads by vortexing.

Transfer the bead suspension to 0.2 ml PCR strip tubes. Spin briefly.

25. Incubate the bead suspensions for 1 min at 95°C in thermal cycler with heated lid.

Immediately transfer the PCR strip tubes to a 96-well magnetic rack.

Transfer the **supernatant**, which contains the library molecules, to a fresh 0.5 ml tube.

qPCR to determine cycle number

Dilute samples 1/40

N	Sybr MM	IS4	Tag	H2O	DNA	Total
1	12.5 µl	0.5 µl	0.5 µl	10.5 µl	1 µl	25 µl

95°C 10mins → 95°C 30s → 60°C 30s → 72°C 30s

45 cycles

Index PCR

To add index to and amplify SSlibs.

N	IS4	Tag	Taq Gold	dNTPs	MgCl2	10X	H2O	Betaine	DNA	Total
1	2 µl	2 µl	2 µl	2 µl	8.0 µl	10 µl	44 µl	20 µl	10 µl	100 µl

95°C 10mins → 95°C 30s → 60°C 30s → 7 °C 30s

x cycles

Number of cycles determined from qPCR

Following index PCR, clean up using Qiagen PCR purification kit following manufacturers protocol.

Overamplification of libraries for capture

10 µl	5x Kapa Buffer
1.5 µl	dNTPs (10mM)
1.5 µl	Sol_amp_p5
1.5 µl	Sol_amp_p7
1 µl	HiFiEnzyme
33.5 µl	ddH2O

49 µl of the above master mix to 1 µl of template (i.e. Clean immort. library)

94ºC	5 min	
94ºC	20s	x 18
55ºC	55s	
72ºC	15s	
72ºC	5 min	
10ºC	forever	

Following library amplification, perform clean-up of PCR products using manufacturers protocol. Quantify using preferred method e.g. Qbit, BioAnalyser etc.

For capture: MYcroarray MYbaits protocol.

http://www.mycroarray.com/mybaits/manuals.html

Post Capture PCR

2x 15 µl DNA in 100 µl volume PCR 16 cycles

N	Primer IS5(10uM)	Primer IS6(10uM)	Hifi Enzyme	dNTPs	5X Kapa buffer	H2O	DNA	Total
1	3 µl	3 µl	2 µl	3 µl	20 µl	54 µl	15 µl	100 µl

94ºC	5 min	
94ºC	20s	
55ºC	55s	x16
72ºC	15s	
72ºC	5 min	
10ºC	forever	

Following Post Capture PCR purification, quantify, and pool for sequencing.

Appendix 2- Published Papers

An 'Aukward' Tale: A Genetic Approach to Discover the Whereabouts of the Last Great Auks

Jessica E. Thomas [1,2,*], Gary R. Carvalho [1,†], James Haile [2,†], Michael D. Martin [3], Jose A. Samaniego Castruita [2], Jonas Niemann [2], Mikkel-Holger S. Sinding [2,4], Marcela Sandoval-Velasco [2], Nicolas J. Rawlence [5], Errol Fuller [6], Jon Fjeldså [7], Michael Hofreiter [8], John R. Stewart [9], M. Thomas P. Gilbert [2,3,†] and Michael Knapp [10,†]

Genes **2017**, *8*(6), 164; doi:10.3390/genes8060164
Academic Editor: J. Peter W. Young
Received: 30 May 2017 / Accepted: 9 June 2017 / Published: 15 June 2017

Genes **2017**, *8*, 164

genes

MDPI

Article

An 'Aukward' Tale: A Genetic Approach to Discover the Whereabouts of the Last Great Auks

Jessica E. Thomas [1,2,*], Gary R. Carvalho [1,†], James Haile [2,†], Michael D. Martin [3], Jose A. Samaniego Castruita [2], Jonas Niemann [2], Mikkel-Holger S. Sinding [2,4], Marcela Sandoval-Velasco [2], Nicolas J. Rawlence [5], Errol Fuller [6], Jon Fjeldså [7], Michael Hofreiter [8], John R. Stewart [9], M. Thomas P. Gilbert [2,3,†] and Michael Knapp [10,†]

1 Molecular Ecology and Fisheries Genetics Laboratory, School of Biological Sciences, Bangor University, Bangor, Gwynedd LL57 2UW, UK; g.r.carvalho@bangor.ac.uk
2 Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5–7, 1350 Copenhagen K, Denmark; drjameshaile@gmail.com (J.H.); jose.samaniego@snm.ku.dk (J.A.S.C.); j.niemann@snm.ku.dk (J.N.); mikkel.sinding@snm.ku.dk (M.-H.S.S.); marcela.velasco@snm.ku.dk (M.S.-V.); mtpgilbert@gmail.com (M.T.P.G.)
3 Department of Natural History, Norwegian University of Science and Technology, University Museum, NO-7491 Trondheim, Norway; mike.martin@ntnu.no
4 Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, N-0318 Oslo, Norway
5 Otago Palaeogenetics Laboratory, Department of Zoology, University of Otago, Dunedin 9054, New Zealand; nic.rawlence@otago.ac.nz
6 65 Springfield Road, Southborough, Tunbridge Wells TN4 0RD, Kent, UK; errolfuller123@btinternet.com
7 Center for Macroecology, Evolution and Climate, the Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark; jfjeldsaa@snm.ku.dk
8 Department of Mathematics and Natural Sciences, Evolutionary Adaptive Genomics, Institute for Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany; michi@palaeo.eu
9 Faculty of Science and Technology, Bournemouth University, Dorset BH12 5BB, UK; jstewart@bournemouth.ac.uk
10 Department of Anatomy, University of Otago, 270 Great King Street, Dunedin 9016, New Zealand; michael.knapp@otago.ac.nz
* Correspondence: bsp20a@bangor.ac.uk
† These authors contributed equally to this work.

Academic Editor: J. Peter W. Young
Received: 30 May 2017; Accepted: 9 June 2017; Published: 15 June 2017

Abstract: One hundred and seventy-three years ago, the last two Great Auks, *Pinguinus impennis*, ever reliably seen were killed. Their internal organs can be found in the collections of the Natural History Museum of Denmark, but the location of their skins has remained a mystery. In 1999, Great Auk expert Errol Fuller proposed a list of five potential candidate skins in museums around the world. Here we take a palaeogenomic approach to test which—if any—of Fuller's candidate skins likely belong to either of the two birds. Using mitochondrial genomes from the five candidate birds (housed in museums in Bremen, Brussels, Kiel, Los Angeles, and Oldenburg) and the organs of the last two known individuals, we partially solve the mystery that has been on Great Auk scholars' minds for generations and make new suggestions as to the whereabouts of the still-missing skin from these two birds.

Keywords: ancient DNA; extinct birds; mitochondrial genome; museum specimens; palaeogenomics

1. Introduction

Over the past three decades, the field of ancient DNA (aDNA) has grown considerably, from sequencing a small section of mitochondrial DNA from the Quagga, an extinct form of the plains zebra [1], to whole genome sequencing from samples up to 735,000 years old [2]. Ancient DNA has been used to answer and address a diverse range of ecological and evolutionary questions, providing insight into countless species' pasts, including our own. However, aDNA can also be a useful tool for museums, specifically for species identification and, under suitable circumstances for reconstructing the history of specimens where museum records are insufficient. This study traces the whereabouts of the skins from the last two documented Great Auks using a palaeogenomic approach.

The Great Auk (Figure 1), *Pinguinus impennis*, Bonnaterre (1790) (traditionally *Alca impennis*, Linnaeus, 1758), has been described as "*perhaps the most curious of all vanished birds*" [3]. It was a bird whose life and ultimate extinction has generated ongoing interest, with several scholars dedicating their lives to Great Auk research [3–7]. Even now, 173 years after the death of the last two recorded captured individuals, there are still many unanswered questions concerning aspect of its life-history, evolution, and extinction. One such mystery that surrounds the Great Auk is the whereabouts of the skins from the last documented pair. In order to be able to correlate the phenotype of the last birds with genomic information obtained from the well-preserved organs, and in view of the active role that researchers and research institutions played in pushing the Great Auk towards extinction, it is of relevance to be able to trace these skins.

Figure 1. A mounted Great Auk skin, The Brussels Auk (RBINS 5355) (MK135), from the collections at Royal Belgian Institute of Natural Sciences (Credit Thierry Hubin (RBINS)).

Once found in great numbers across the North Atlantic (Figure 2), this flightless bird was heavily hunted for its meat, oil, and feathers. By the start of the 19th century, populations in the North-West Atlantic had been decimated. The last few remaining birds were breeding on the skerries off the south-west coast of Iceland, but with their scarcity increasing, Great Auks were then also sought after as a desirable item for both private and institutional collections [3,5,8–10].

From 1830 to 1841, several trips were taken to Eldey Island (Figure 2) where Great Auks were caught, killed, and sold for exhibitions. Following a three-year period of no recorded captures of Great

Auks, Carl Siemsen commissioned an expedition to Eldey to search for any remaining birds. Between 2 and 5 June 1844, the expedition reached Eldey Island where two Great Auks were observed amongst smaller birds inhabiting the island. Both Auks were killed and their broken egg discarded. The birds, though, were never to reach Siemsen. The expedition leader sold them to Christian Hansen, who then sold them to the apothecary Möller, in Reykjavik, Iceland. Möller skinned the birds and sent them, as well as their preserved body parts, to Denmark [3,6,7].

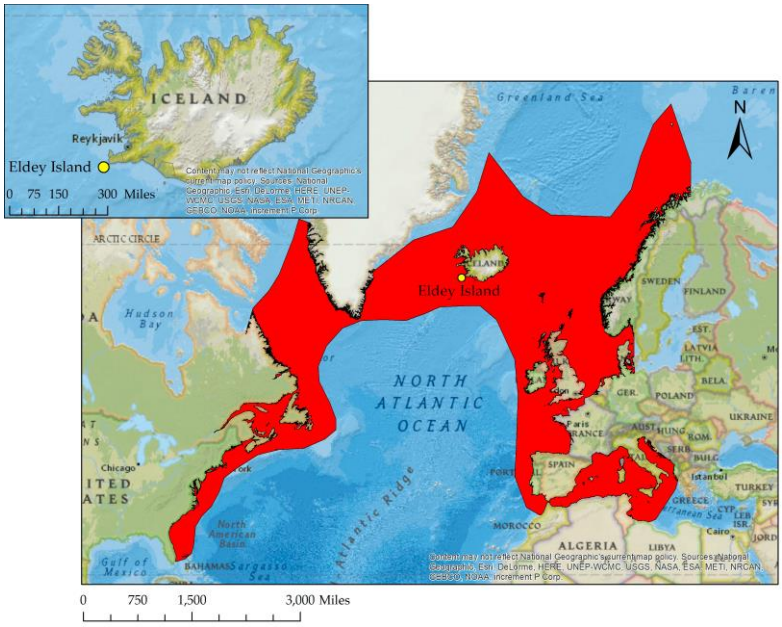

Figure 2. The Great Auk's breeding range across the North Atlantic, as indicated by the red area and the location of Eldey Island (yellow dot) off the south-west coast of Iceland, the site where the last documented Great Auks were killed. Maps were created using spatial data provided by BirdLife International/IUCN [11] with the National Geographic basemap in ArcGIS (ESRI, Redlands, CA, USA) [12].

The internal organs of these two birds now reside in the Natural History Museum of Denmark. However, the location of the skins of those individuals remains a mystery, despite considerable effort of notable Great Auk scholars to solve it.

Fuller [3] describes in detail the known history of the 80 or so specimens that are still in existence in collections today and concludes: "*Somehow, amid all the frantic Garefowl* [another name for Great Auk] *research of the nineteenth century, they* [the skins] *were lost track of. Several of the surviving stuffed specimens, notably those in Kiel, Bremen and Oldenburg were tentatively identified with them. The most likely candidates, however, are the birds now in Los Angeles and in Brussels*" [3] (p. 85).

Our study compares complete mitochondrial genome (mitogenome) sequences from the five candidate skins (those housed in Bremen, Brussels, Kiel, Los Angeles, and Oldenburg) to the internal organs of the last documented captured Great Auks (stored in Copenhagen) to test which—if any—of Fuller's candidate skins likely belong to one of the last two individuals.

1. Materials and Methods

1.1. Sample Information

Specimens from the candidate list proposed by Fuller [3] and the organs from the two 1844 Eldey Island individuals, were sampled using sterile equipment and the appropriate method for sample type, which caused minimal physical damage to the specimen (Table 1).

Table 1. Sample information. Lab ID number used during laboratory and analysis process. Mount name and description given by Fuller and its number in various published lists of Great Auk mounts [3]. Origin and date information as noted by Fuller [3]. Institution information relating to the present location of specimen and the curator/sample collector name.

Lab ID	Bird Name, Number & Description	Origin & Date	Institution	Curator/Collector	Institution Number	Sample Type/Sampling Method
MK131	Last Great Auk 1 Oesophagus (male)	Eldey Island, Iceland. Date: June 1844	Natural History Museum of Denmark. Copenhagen, Denmark	J. Fjeldså/ J. Thomas	NHMD 153069	Oesophagus. Tissue cut from end of oesophagus.
MK132	Last Great Auk 2 Oesophagus (female)	Eldey Island, Iceland. Date: June 1844	Natural History Museum of Denmark. Copenhagen, Denmark	J. Fjeldså/ J. Thomas	NHMD 153070	Oesophagus. Tissue cut from end of oesophagus.
MK133	The Oldenburg Auk Fuller: Bird no. 47, Grieve: no. 57, Hahn: no. 77 Adult in summer plumage	Iceland. Probably Eldey. Date: Unknown	Landesmuseum Natur und Mensch Oldenburg. Germany	C. Barilaro	AVE 8086	Body tissue. Tissue cut from body of bird under wing.
MK134	The Bremen Auk Fuller: Bird no. 36, Grieve: no. 10, Hahn: no. 71 Adult in summer plumage	Unknown. Probably Eldey. Date: Unknown	Übersee-Museum Bremen. Germany	M. Stiller	RKNr. 2392	Toepad tissue. Tissue cut from feet.
MK135	The Brussels Auk Fuller: Bird no. 3, Grieve: no. 15, Hahn: no. 6 Adult in summer plumage	Probably Eldey Date: Unknown perhaps June, 1844	Institut Royal des Sciences Naturelles de Belgique. Brussels, Belgium	G. Lenglet	RBINS 5355	Toepad tissue. Tissue cut from feet
MK136	Dawson Rowley's Los Angeles Auk Fuller: Bird no. 73, Grieve no. 13,	Iceland. Probably Eldey.	Natural History Museum of Los	K. Garett	LACM	Feather.
	Hahn: no. 5 Adult in summer plumage, said to be female	Date: Unknown perhaps June, 1844	Angeles County. USA		76476	Feathers plucked from body of bird.
MK138	The Schleswig-Holstein Auk Fuller: Bird no. 42, Grieve: no. 31, Hahn: no. 74 Adult in summer plumage	Unknown Date: Unknown	Zoologisches Museum der Christian-Albrechts Universität zu Kiel. Germany	D. Brandis/ L. Rosotta	cat. No. A0585	Toepad tissue. Tissue cut from feet.
LastGA2_Heart	Last Great Auk 2 Heart (female)	Eldey Island, Iceland. Date: June 1844	Natural History Museum of Denmark. Copenhagen, Denmark	J. Fjeldså/ J. Haile	NHMD 153070	Heart. Tissue cut from aorta.

1.1. DNA Extraction

All lab work prior to polymerase chain reaction (PCR) amplification was carried out in designated aDNA laboratories that adhere to strict aDNA protocols [13]. For each DNA extraction and library build, negative controls were used to check for contamination by exogenous DNA. All post-PCR work on amplified DNA was carried out in separate laboratory facilities.

Genomic DNA was extracted from the oesophagus (Figure 3a), skin (Figure 3b), toepad tissue (Figure 3c), and feathers using a modified version of Dabney et al. [14] in which the initial digestion was carried out following the protocol by Gilbert et al. [15]. This digestion buffer is better suited to extraction from these tissues types than the Dabney et al. [14] digestion buffer, which was optimised for DNA extraction from bone. Subsequent DNA purification and elution was conducted following the approach described by Dabney et al. [14]. Genomic DNA was extracted from the heart tissue (Figure 3d) using the protocol by Campos et al. [16].

Figure 3. (**a**) Jars containing the oesophagus from the last two individuals killed on Eldey Island. The oesophagus from the larger jar represents that of the individual labelled male (NHMD153069) (MK131). The smaller jar contains the oesophagus from the female bird (NHMD153070) (MK132) (credit. J. Thomas). (**b**) Sampling of The Oldenburg Auk (AVE 8086) (MK133) to remove a section of body tissue for DNA extraction (credit. C. Barilaro, Landesmuseum Natur und Mensch Oldenburg). (**c**) Sampling the toe pad of The Bremen Auk (RKNr. 2392) (MK134) to remove tissue sample (credit M. Stiller, Übersee-Museum Bremen). (**d**) The hearts from the last two documented individuals. The heart from the female individual has been sampled for this study (top) (NHMD153070) (LastGA2_Heart) (credit Natural History Museum of Denmark).

1.1. Data Generation

Single stranded libraries were constructed for all samples, except LastGA2_Heart, following Gansauge & Meyer [17], with modifications as described by Bennett et al. [18], as this allowed for targeting of the smallest fragments of DNA, typical of highly degraded specimens. For LastGA2_Heart, the protocol described by Meyer & Kircher [19] was used. Enrichment for complete mitogenomes was performed using MYcroarray MYbaits, following the manufacturer's manual v2.3.1 [20] on all samples except MK138 and LastGA2_Heart. Samples were sequenced on Illumina platforms (HiSeq and MiSeq) by New Zealand Genomics Limited, Otago Branch, or the Danish National High-Throughput DNA Sequencing Centre.

1.2. Read Processing

Processing of raw sequence data was facilitated by the PALEOMIX v1.2.5 pipeline [21], which performs adapter trimming, read mapping to a reference genome, and quality-based filtering. Low-quality bases and adapter sequences were trimmed from the 3' ends of DNA reads with the software AdapterRemoval v2.1.7 [22,23] using a mismatch rate of 0.333 (command-line option—mm 3). Paired end reads overlapping by at least 11 base pairs (bp) were collapsed into a single read with re-calibrated base quality scores. Trimmed reads shorter than 25 bp were discarded.

Mapping to the Great Auk reference mitogenome (GenBank: KU158188.1) [24] was performed with Burrows–Wheeler Aligner (BWA) v0.5.10 [25] with seeding deactivated and otherwise default settings. PCR duplicates were removed with the MarkDuplicates function within Picard v1.82 [26] and the rmdup function within the software SAMtools [27]. Collapsed reads were filtered using a script included with PALEOMIX. Reads with mapping quality (MAPQ) scores <20 were removed from further analysis. Local realignment of reads misaligned to the reference mitogenome was performed with the RealignerTargetCreator and IndelRealigner tools included in the software Genome Analysis Toolkit (GATK) v3.6.0 [28]. The pipeline also utilised MapDamage2 [29] to recalibrate base qualities of aligned sequence reads in each sequencing library in order to remove the residual aDNA damage patterns. The UnifiedGenotyper algorithm within GATK v3.6.0 was used to determine haploid genotypes within individual samples.

A relaxed and strict filtering system was used to create consensus sequences and alignments from the processed data. In the first stage of filtering, both systems used VCFtools [30] to filter genotypes from the final alignment when their genotype quality scores were less than 30. For the relaxed alignment, the per-individual read depth was set to only include bases with a minimum of 3-fold coverage. Bases called for the consensus sequence had to be present at a frequency higher than 33%. To be included in the final alignment, no more than 33% of bases could be missing from the consensus sequence of an individual.

For the strict settings, the per-individual read depth was set to only include bases with at least 10-fold coverage. Geneious v-10.1.3 [31] was used to filter bases so that the majority base was present in more than 90% of reads. For an individual to be included in the final alignment, no more than 20% of sites could be missing from the individual's consensus sequence.

A custom script was used to convert the filtered Variant Call Format (VCF) file into a multiple sequence alignment in FASTA format.

Following read processing, the data was aligned using Seaview v4.0 [32] with the algorithm *Muscle -maxiters2 -diags*. The alignment was manually checked for errors using BioEdit v7.2.5 [33], and Tablet v-1.16.09.06 [34] was used to view the rescaled Binary Alignment Map (BAM) file for each sample.

MEGA v-7.0.21 [35] was used to generate a pairwise distance table for all sequenced individuals. Phylogenetic relationships between the individuals were reconstructed and visualized using a maximum-likelihood approach as implemented in MEGA v-7.0.21 [35]. jModelTest v-2.1.10 [36,37] was used to determine the most suitable nucleotide substitution model, which was a Hasegawa–Kishino–Yano (HKY) [38] model. Initial trees for the heuristic search were obtained by

applying Neighbour-Joining methods to a matrix of pairwise distances estimated using the maximum composite likelihood approach. Branch lengths are measured in number of substitutions per site. All positions containing gaps and missing data were removed. Phylogenies were reconstructed from 500 bootstrap pseudoreplicates to evaluate branch support.

1. Results

Mitogenome sequence data was obtained from all candidate specimens as well as from the two oesophagi of the last Great Auks. Unique coverage of the mitogenomes for these samples ranged from 6.2× to 288.6× (Table 2). As DNA extracted from the oesophagus of the female last Great Auk (MK132) yielded only a low coverage, poor quality mitogenome assembly, DNA from the heart of the same individual was also sequenced. This yielded a high coverage (430×) mitogenome, which was used in all further analyses.

Table 2. Read processing results for all samples.

Sample	GenBank Accession Number	Number of Reads	Number of Unique Reads Mapping to Reference Mitogenome	Estimated Coverage from Unique Hits	Relaxed Settings Sequence Length (bp[1])	Strict Settings Sequence Length (bp)
MK131	MF188883	300754	30,297	74.40	16,001	15,067
MK132	NA	550631	2366	6.23	13,267	3312
MK133	MF188884	429392	8750	23.04	16,251	14,240
MK134	MF188885	343766	86,325	288.62	16,607	16,526
MK135	MF188886	579992	27,767	88.90	16,554	16,356
MK136	MF188887	563635	24,401	67.83	16,330	15,833
MK138	MF188888	10796460	2799	9.76	16509	7866
LastGA2_Heart	MF188889	957970612	121,886	430.09	16,698	16,649

With the sequence data from the heart of the female last Great Auk (LastGA2_Heart), the alignment of all sequences assembled under the relaxed rules had a length of 15,790 bp after sites not covered by all consensus sequences were removed. For the strict alignment, MK138 did not meet criteria set by the strict filtering settings as more than 20% sites were missing. With this individual removed, we obtained a strict alignment length of 13,475 bp.

The pairwise distance matrix (Table 3) shows that the consensus sequence obtained from sample MK131, the oesophagus of the male, is identical to the consensus sequence obtained from MK135, The Brussels Auk. No other consensus sequences match. LastGA2_Heart, the female last Great Auk, groups with MK136 and MK134 in the maximum likelihood phylogeny (Figure 4), but there are 18 and 20 well-supported differences between the consensus sequences, respectively. Analysis presented here was generated using data from the relaxed filtering settings, but results were consistent with data from the strict filtering system. Thus, only the male last Great Auk has a corresponding DNA match among the candidate skin samples identified by Fuller [3].

Table 3. Pairwise distance matrix. Estimates of evolutionary divergence between sequences generated using the relaxed settings. The number of base differences per sequence from between sequences are shown. All positions containing gaps and missing data were removed, leaving a total of 15,790 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [35].

	MK131	MK133	MK134	MK135	MK136	MK138	LastGA2_Heart
MK131_LastGA1							
MK133_Oldenburg	17						
MK134_Bremen	18	23					
MK135_Brussels	0	17	18				
MK136_LA	16	23	20	16			
MK138_Kiel	14	11	20	14	20		
LastGA2_Heart	16	23	20	16	18	20	

Figure 4. Maximum liklihood reconstruction of phylogenetic relationships between individuals, under the relaxed filtering settings. Branch labels are bootstrap support values for the respective sample. Evolutionary analyses were conducted in MEGA7 [35].

1. Discussion

The genetic analyses presented here help to partially resolve the mystery of the missing skins of the last two Great Auks. They provide evidence of matching mitochondrial genomes for the internal organs of the last male Great Auk held at the Natural History Museum of Denmark in Copenhagen and the Great Auk skin held at the Royal Belgian Institute of Natural Sciences, Brussels (Figure 1). Mitochondrial DNA cannot always be unambiguously used in identification of individuals. However, in a broader study of forty one Great Auk mitogenomes from across their range, Thomas et al. (in prep) [39], found that mitochondrial diversity in Great Auks remained high right up to their demise, with no other individuals found to have the same mitochondrial haplotype. Together with the information from the historical record, the match between the internal organs and The Brussels Auk therefore appears to be more than just a coincidence.

There are around 80 known mounted Great Auk skins in museums worldwide. However, the majority can be ruled out of any speculation that they belonged to the last pair due to their history (for example, if they were collected before 1844). Those tested in the current study were placed on the candidate specimen list due to several factors that led Fuller, as well as other experts like the University of Copenhagen Professor Japetus Steenstrup (dubbed 'Father of Garefowl History' by Grieve, 1885), and Grieve [4], to suspect that they originated from the 1844 Eldey pair. Details such as when and where they were acquired, from whom (i.e., the dealer), and suggestions by renowned Great Auk scholars made the birds in Bremen, Brussels, Kiel, Los Angeles, and Oldenburg the top candidates [3].

In the museum industry, accurate records and archiving are obviously of high priority, with labels and registers providing vital information about the specimens [40–42]; it therefore seems unexpected that the two bird skins could have been "lost". However, at the time, their significance as the final remnants of the species was not recognised. The story of the ending of these individuals lives is well documented due to the efforts of English naturalist John Wolley and Cambridge University Professor Alfred Newton, who travelled to Iceland in the late 1850s and spoke directly with those who were part of the 1844 Eldey Island voyage (details from Wolley's notebook 'Garefowl books' published in Newton, 1861 [7]). What happened once the skins and their organs reached Denmark, on the other hand, is poorly recorded and remains speculative [3].

In the archives of Cambridge University are the fragments of information that Newton learned of the birds. On notes dated 1861, it was recorded that Professor Reinhardt of the Royal Museum (Copenhagen) believed the skins and their organs had been purchased for the museum by Professor Eschricht of the University of Copenhagen. He is said to have taken the skins to the Congress of German Naturalists in Bremen in the autumn of 1844 [3].

The connection with the skins and the Congress in Bremen could be what led Steenstrup to inform Grieve of his suspicions that the specimen at the museum in Bremen (MK134) was indeed one of the last birds [4]. Yet, this bird was bought by the museum at the time of the Congress from the Hamburg dealer Salmin, not Eschricht. Therefore, while the possibility may be there for Salmin to have first had the bird from Eschricht and then sold it on, it is also likely that it was a bird he had in his stocks prior to 1844 [3]. This study shows The Bremen Auk is not a match with either of the organs from the last pair, suggesting that it did indeed come from an earlier raid of Eldey.

The specimen in Kiel, the Schleswig–Holstein Auk (MK138), was purchased in 1844. With such a suggestive purchase date it is a contender in the mystery [3]. Professor Steenstrup was quoted by Grieve as saying, "*If really purchased in 1844, it might perhaps be the second of these two Garefowls got in 1844, but traditionally I never heard that mentioned*" [4] (Grieve Appendix p. 13 [4]). Our study shows this specimen was not a match, so Steenstrup was correct in his belief.

With regard to The Oldenburg Auk (MK133), this specimen was once regarded by nineteenth century scholars as belonging to one of the last birds. However, the records for this bird shows it was obtained prior to 1844 and is therefore ruled out [3]. It was tested in this study due to the suggestions of these early researchers but was not a match.

The history of The Brussels Auk (MK135) and Dawson Rowley's Los Angeles Auk (MK136) can be traced back to 1845 when they were said to be in the hands of a well-known, and well connected, Great Auk dealer, Israel of Copenhagen. Israel is known to have had excellent links with Iceland and spent his winters in Copenhagen and his summers in Amsterdam [3]. Fuller suggests that perhaps Israel, if he did not receive them direct from Iceland, purchased the birds in Bremen from Eschricht. The birds have a detailed history, passing through the hands of several dealers. From Israel, they were bought by Lintz, a Hamburg merchant, and in 1845 were sold on to the Amsterdam branch of the dealer, Frank. In Newton's notes at Cambridge it was recorded that Frank believed the two skins he bought were from the last pair. The Brussels Auk was purchased in 1847 by Viscount Bernard Du Bus Ghisignies, director of the Brussels Museum [3]. The history of The Brussels Auk therefore strongly supports our positive match with MK131.

If the bird in Brussels, which came from Israel of Copenhagen, is from one of the last birds, then this would suggest that the second bird he had would also be from Eldey in 1844 and therefore be a positive match with the second set of organs. Israel's second bird has an even longer story than that of MK135, but it now resides in the Natural History Museum of Los Angeles County [3]. This specimen, Dawson Rowley's Los Angeles Auk (MK136), was tested, and the results showed it did not match LastGA2_Heart. With this negative result, we can only speculate which of the remaining untested birds could be identified as the second individual.

A possible scenario to explain the mismatch between Dawson Rowley's Los Angeles Auk (MK136) and the internal organs from the Natural History Museum of Denmark involves a mix up of skins. Dawson Rowley's Los Angeles Auk, was once one of two Great Auks owned by George Dawson Rowley. During the 1930s, they were passed to Captain Vivian Hewitt who owned two additional specimens. The four specimens are currently held in Cardiff, Birmingham, Los Angeles, and Cincinnati. At Hewitt's death, his collection had been put under the control of Spink and Son Ltd., a London dealer, who offered them for sale. While organising Hewitt's affairs, the four birds were mixed up. The identity of the birds now in Birmingham and Cardiff could be easily resolved, but those now in Los Angeles and Cincinnati are harder to determine. It is thought that their identities could be determined from annotated photographs taken in 1871 by George Dawson Rowley when they were in his possession [3]. However, we speculate that their identities were not correctly resolved and that perhaps the bird in Cincinnati was the original bird from Israel of Copenhagen. If this were the case, then it would explain why the Los Angeles bird fails to match with either of the last Great Auk organs held in Copenhagen.

In summary, we suggest that The Brussels Auk is the skin from the last male Great Auk killed on Eldey Island in June 1844. The skin of the female killed at the same time remains unaccounted for, but a common history with The Brussels Auk makes the skin currently held at Cincinnati Museum of Natural History and Science, a likely candidate. A re-evaluation of the historical records may reveal further candidate skins amongst those currently held in museums around the world.

1. Conclusions

Ancient DNA has been used to evaluate museum collections in the past, albeit usually for taxonomic identification of unidentified or misidentified accessions. Our study shows an alternative use of the technology. It demonstrates the utility of molecular tools and advanced sequencing to contribute to questions, which are not primarily biological or molecular but rather historical in nature. The unraveling of the mystery surrounding the whereabouts of the skins of the last two Great Auks represents a fascinating element in the story of extinction and human involvement in that process.

Acknowledgments: We are very grateful to the institutions that provided samples to this project and the curators/sample collectors within them (Jon Fjeldså at Natural History Museum of Denmark, Christina Barilao at Landesmuseum Natur und Mensch Oldenburg, Germany, Michael Stiller at Übersee-Museum Bremen, Germany, Georges Lenglet at Institut Royal des Sciences Naturelles de Belgique, Brussels, Belgium (RBINS), Kimball Garret at Los Angeles County Museum of Natural History, Los Angeles, USA and Dirk Brandis and Lina Rosotta at Zoologisches Museum der Christian-Albrechts-Universität zu Kiel, Germany). We thank Jan Bolding Kristensen for his help and advice with the sampling of the Great Auk organs and photograph of the Great Auk hearts. We also thank Thierry Hubin (RBINS), Michael Stiller (Kiel), and Christina Barilaro for providing photographs for use in this publication. Funding was provided through NERC Ph.D. Studentship (NE/L501694/1), the Genetics Society-Heredity Fieldwork Grant, and European Society for Evolutionary Biology–Godfrey Hewitt Mobility Award. MK is supported by a Rutherford Discovery Fellowship from the Royal Society of New Zealand. We thank members of the Molecular Ecology and Fisheries Genetics Laboratory at Bangor University, EvoGenomics & GeoGenetics at University of Copenhagen, and the Biological Anthropology group at the University of Otago, for guidance in the laboratory, analysis, and useful discussions. Sequencing was provided by either The Danish National High-Throughput DNA Sequencing Centre or New Zealand Genomics Limited.

Author Contributions: J.E.T. conceived the study, J.E.T., J.H., G.R.C., M.T.P.G., and M.K. designed the experiments; J.E.T., J.H., M-H.S.S., and M.S.-V. conducted the experiments; J.E.T., M.D.M., J.A.S.C., and J.N. analysed the data; N.J.R., M.H., and J.R.S. provided the initial framework for the study; all authors contributed to writing the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Higuchi, R.; Bowman, B.; Freiberger, M.; Ryder, O.A.; Wilson, A.C. DNA sequences from the quagga, an extinct member of the horse family. *Nature* **1984**, *312*, 282–284. [CrossRef]
2. Orlando, L.; Ginolhac, A.; Zhang, G.; Froese, D.; Albrechtsen, A.; Stiller, M.; Schubert, M.; Cappellini, E.; Petersen, B.; Moltke, I.; et al. Recalibrating equus evolution using the genome sequence of an early middle pleistocene horse. *Nature* **2013**, *499*, 74–78. [CrossRef]
3. Fuller, E. *The Great Auk*; Errol Fuller: Kent, UK, 1999; 448p; ISBN 0-9533553-0-6.
4. Grieve, S. *The Great Auk, or Garefowl. Its History, Archaeology and Remains*, Digitally Printed Version 2015 ed.; Cambridge University Press: Cambridge, UK, 1885; ISBN 978-1-108-08147-4.
5. Bengtson, S.-A. Breeding ecology and extinction of the great auk (*Pinguinus impennis*): Anecdotal evidence and conjectures. *Auk* **1984**, *101*, 1–12.
6. Gaskell, J. *Who Killed the Great Auk?* Oxford University Press: New York, NY, USA, 2000; 240p; ISBN 978-0-19856478-2.
7. Newton, A. XIII.—Abstract of Mr. J. Wolley's researches in Iceland respecting the gare-fowl or great auk (*Alca impennis*, linn.). *Ibis* **1861**, *3*, 374–399. [CrossRef]
8. Meldgaard, M. The great auk, *Pinguinus impennis* (L.) in greenland. *Hist. Biol.* **1988**, *1*, 145–178. [CrossRef]
9. Serjeantson, D. The great auk and the gannet: A prehistoric perspective on the extinction of the great auk. *Int. J. Osteoarchaeol.* **2001**, *11*, 43–55. [CrossRef]
10. Montevecchi, W.A.; David, A. *Kirk. Great Auk (Pinguinus impennis) Birds of North America Online*; Rodewald, P.G., Ed.; Cornell Lab of Ornithology: Ithaca, NY, USA, 1996.
11. International Union for Conservation of Nature, Birdlife International and Handbook of the Birds of the World (2016). *Pinguinus impennis*, The IUCN Red List of Threatened Species. Version 2016–*3*; 2016. Available online: http://maps.iucnredlist.org/map.html?id=22694856 (accessed on 16 May 2017).
12. ArcGIS. *Arcgis Desktop, Arcmap*; 10.5.0.6491; ESRI: Redlands, CA, USA, 2016.
13. Knapp, M.; Clarke, A.C.; Horsburgh, K.A.; Matisoo-Smith, E.A. Setting the stage—Building and working in an ancient DNA laboratory. *Ann. Anat.* **2012**, *194*, 3–6. [CrossRef] [PubMed]
14. Dabney, J.; Knapp, M.; Glocke, I.; Gansauge, M.-T.; Weihmann, A.; Nickel, B.; Valdiosera, C.; García, N.; Pääbo, S.; Arsuaga, J.-L.; et al. Complete mitochondrial genome sequence of a middle pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 15758–15763. [CrossRef]
15. Gilbert, M.T.P.; Tomsho, L.P.; Rendulic, S.; Packard, M.; Drautz, D.I.; Sher, A.; Tikhonov, A.; Dalén, L.; Kuznetsova, T.; Kosintsev, P.; et al. Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. *Science* **2007**, *317*, 1927–1930. [CrossRef] [PubMed]
16. Campos, P.F.; Gilbert, T.M.P. DNA extraction from keratin and chitin. In *Ancient DNA: Methods and Protocols*; Shapiro, B., Hofreiter, M., Eds.; Humana Press: Totowa, NJ, USA, 2012; pp. 43–49; ISBN 978-1-61779-516-9.
17. Gansauge, M.-T.; Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **2013**, *8*, 737–748. [CrossRef]
18. Bennett, E.A.; Massilani, D.; Lizzo, G.; Daligault, J.; Geigl, E.M.; Grange, T. Library construction for ancient genomics: Single strand or double strand? *BioTechniques* **2014**, *56*, 289–290, 292–286, 298, passim. [CrossRef]
19. Meyer, M.; Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, *2010*. [CrossRef]
20. MYcroarry. *Mybaits Manual – Sequence Enrichment for Targeted Sequencing v2.3.1*; MYcroarry, 2014; Available online: http://www.mycroarray.com/mybaits/manuals.html (accessed on 13 June 2017).
21. Schubert, M.; Ermini, L.; Sarkissian, C.D.; Jónsson, H.; Ginolhac, A.; Schaefer, R.; Martin, M.D.; Fernández, R.; Kircher, M.; McCue, M.; et al. Characterization of ancient and modern genomes by snp detection and phylogenomic and metagenomic analysis using paleomix. *Nat. Protoc.* **2014**, *9*, 1056–1082. [CrossRef]
22. Lindgreen, S. Adapterremoval: Easy cleaning of next-generation sequencing reads. *BMC Res. Notes* **2012**, *5*, 337. [CrossRef]
23. Schubert, M.; Lindgreen, S.; Orlando, L. Adapterremoval v2: Rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **2016**, *9*, 88. [CrossRef]
24. Anmarkrud, J.A.; Lifjeld, J.T. Complete mitochondrial genomes of eleven extinct or possibly extinct bird species. *Mol. Ecol. Resour.* **2017**, *17*, 334–341. [CrossRef]

1. Li, H.; Durbin, R. Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef]
2. BroadInstitute. Picard v1.82. Available online: http://broadinstitute.github.io/picard/ (accessed on 13 June 2017).
3. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and samtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef]
4. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The genome analysis toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [CrossRef]
5. Jónsson, H.; Ginolhac, A.; Schubert, M.; Johnson, P.L.F.; Orlando, L. Mapdamage2.0: Fast approximate bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **2013**, *29*, 1682–1684. [CrossRef]
6. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and vcftools. *Bioinformatics* **2011**, *27*, 2156–2158. [CrossRef]
7. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef]
8. Gouy, M.; Guindon, S.; Gascuel, O. Seaview version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **2010**, *27*, 221–224. [CrossRef]
9. Hall, T.A. Bioedit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/nt. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98.
10. Milne, I.; Stephen, G.; Bayer, M.; Cock, P.J.A.; Pritchard, L.; Cardle, L.; Shaw, P.D.; Marshall, D. Using tablet for visual exploration of second-generation sequencing data. *Brief. Bioinform.* **2013**, *14*, 193–202. [CrossRef]
11. Kumar, S.; Stecher, G.; Tamura, K. Mega7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef]
12. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. Jmodeltest 2: More models, new heuristics and high-performance computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef]
13. Guindon, S.; Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **2003**, *52*, 696–704. [CrossRef]
14. Hasegawa, M.; Kishino, H.; Yano, T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **1985**, *22*, 160–174. [CrossRef]
15. Thomas, J.E.; Haile, J.; Martin, M.D.; Samaniego Castuita, J.A; Niemann, J.; Sinding, .M.-H.S.; Sandoval-Velasco, M..; Soares, A.E.R.; Rawlence, N.J.; Fuller, E.; et al. The evolution and extinction of the Great Auk. (manuscript in preparation).
16. Boessenkool, S.; Star, B.; Scofield, R.P.; Seddon, P.J.; Waters, J.M. Lost in translation or deliberate falsification? Genetic analyses reveal erroneous museum data for historic penguin specimens. *Proc. R. Soc. B Biol. Sci.* **2010**, *277*, 1057–1064. [CrossRef]
17. Rawlence, N.J.; Kennedy, M.; Waters, J.M.; Scofield, R.P. Morphological and ancient DNA analyses reveal inaccurate labels on two of buller's bird specimens. *J. R. Soc. N. Z.* **2014**, *44*, 163–169. [CrossRef]
18. Shepherd, L.D.; Tennyson, A.J.D.; Lambert, D.M. Using ancient DNA to enhance museum collections: A case study of rare kiwi (*Apteryx* spp.) specimens. *J. R. Soc. N. Z.* **2013**, *43*, 119–127. [CrossRef]

References

Adler, C. J., Dobney, K., Weyrich, L. S., Kaidonis, J., Walker, A. W., Haak, W., Bradshaw, C. J. A., Townsend, G., Sołtysiak, A., Alt, K. W., Parkhill, J. and Cooper, A. (2013) 'Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions', *Nature Genetics*. Nature Research, 45(4), pp. 450–455. doi: 10.1038/ng.2536.

Allentoft, M. E., Bunce, M., Scofield, R. P., Hale, M. L. and Holdaway, R. N. (2010) 'Highly skewed sex ratios and biased fossil deposition of moa: ancient DNA provides new insight on New Zealand's extinct megafauna', *Quaternary Science Reviews*. Pergamon, 29(5–6), pp. 753–762. doi: 10.1016/J.QUASCIREV.2009.11.022.

Allentoft, M. E., Collins, M., Harker, D., Haile, J., Oskam, C. L., Hale, M. L., Campos, P. F., Samaniego, J. A., Gilbert, M. T. P., Willerslev, E., Zhang, G., Scofield, R. P., Holdaway, R. N. and Bunce, M. (2012) 'The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils.', *Proceedings. Biological sciences*. The Royal Society, 279(1748), pp. 4724–33. doi: 10.1098/rspb.2012.1745.

Allentoft, M. E., Heller, R., Holdaway, R. N. and Bunce, M. (2015) 'Ancient DNA microsatellite analyses of the extinct New Zealand giant moa (Dinornis robustus) identify relatives within a single fossil site.', *Heredity*. Nature Publishing Group, 115(6), pp. 481–487. doi: 10.1038/hdy.2015.48.

Allentoft, M. E., Heller, R., Oskam, C. L., Lorenzen, E. D., Hale, M. L., Gilbert, M. T. P., Jacomb, C., Holdaway, R. N. and Bunce, M. (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*. National Academy of Sciences, 111(13), pp. 4922–7. doi: 10.1073/pnas.1314972111.

Anijalg, P., Ho, S., Davison, J., Keis, M., Tammeleht, E., Bobowik, K., Tumanov, I., Saveljev, A., Lyapunova, E., Vorobiev, A., Markov, N., Kryukov, A., Kojola, I., Swenson, J., Hagen, S., Eiken, H., Paule, L. and Saarma, U. (2018) 'Large-scale migrations of brown bears in Eurasia and to North America during the Late Pleistocene', *Journal of Biogeography*, In press.

Anmarkrud, J. A. and Lifjeld, J. T. (2017) 'Complete mitochondrial genomes of eleven extinct or possibly extinct bird species', *Molecular Ecology Resources*. Wiley Online Library, 17(2), pp. 334–341. doi: 10.1111/1755-0998.12600.

Ávila-Arcos, M. C., Cappellini, E., Romero-Navarro, J. A., Wales, N., Moreno-Mayar, J. V., Rasmussen, M., Fordyce, S. L., Montiel, R., Vielle-Calzada, J.-P., Willerslev, E. and Gilbert, M. T. P. (2011) 'Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA', *Scientific Reports*. Nature Publishing Group, 1, pp. 1–5. doi: 10.1038/srep00074.

Ávila-Arcos, M. C., Sandoval-Velasco, M., Schroeder, H., Carpenter, M. L., Malaspinas, A.-S., Wales, N., Peñaloza, F., Bustamante, C. D. and Gilbert, M. T. P. (2015) 'Comparative performance of two whole-genome capture methodologies on ancient DNA Illumina libraries', *Methods in Ecology and Evolution*. Edited by M. Bunce, 6(6), pp. 725–734. doi: 10.1111/2041-210X.12353.

Baker, A. J., Pereira, S. L. and Paton, T. A. (2007) 'Phylogenetic relationships and divergence times of Charadriiformes genera: multigene evidence for the Cretaceous origin of at least 14 clades of shorebirds.', *Biology letters*. The Royal Society, 3(2), pp. 205–9. doi: 10.1098/rsbl.2006.0606.

Bandelt, H. J., Forster, P. and Röhl, A. (1999) 'Median-joining networks for inferring intraspecific phylogenies.', *Molecular biology and evolution*, 16(1), pp. 37–48.

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., Marshall, C., McGuire, J. L., Lindsey, E. L., Maguire, K. C., Mersey, B. and Ferrer, E. A. (2011) 'Has the Earth's sixth mass extinction already arrived?', *Nature*, 471(7336), pp. 51–57. doi: 10.1038/nature09678.

Bédard, J. (1969) 'Adaptive radiation in Alcidae', *Ibis*. Blackwell Publishing Ltd, 111(2), pp. 189–198. doi: 10.1111/j.1474-919X.1969.tb02526.x.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W. and Courchamp, F. (2012) 'Impacts of climate change on the future of biodiversity', *Ecology Letters*. Blackwell Publishing Ltd, 15(4), pp. 365–377. doi: 10.1111/j.1461-0248.2011.01736.x.

Bengtson, S. . (1984) 'Breeding ecology and extinction of the Great Auk (*Pinguinus impennis*): Anecdotal evidence and conjectures', *The Auk*. Ornithological Society of North America, 101(1), pp. 1–12.

Bennett, E. A., Massilani, D., Lizzo, G., Daligault, J., Geigl, E. M. and Grange, T. (2014) 'Library construction for ancient genomics: Single strand or double strand?', *BioTechniques*, 56(6), pp. 289–300. doi: 10.2144/000114176.

Bergmann, C. (1847) 'Über die Verhältnisse der Warnoekonomie der Thiere zu ihrer Grosse', *Gottinger Stud.*, 1, pp. 595–708.

Bernardi, G., Wiley, E. O., Mansour, H., Miller, M. R., Orti, G., Haussler, D., O'Brien, S. J., Ryder, O. A. and Venkatesh, B. (2012) 'The fishes of Genome 10K', *Marine Genomics*, 7, pp. 3–6. doi: 10.1016/j.margen.2012.02.002.

Binladen, J., Wiuf, C., Gilbert, M. T. P., Bunce, M., Barnett, R., Larson, G., Greenwood, A. D., Haile, J., Ho, S. Y. W., Hansen, A. J. and Willerslev, E. (2006) 'Assessing the Fidelity of Ancient DNA Sequences Amplified From Nuclear Genes', *Genetics*, 172(2).

BirdLife International (2016a) *Alca torda., The IUCN Red List of Threatened Species 2016*. IUCN Global Species Programme Red List Unit. Available at: http://www.iucnredlist.org/details/22694852/0 (Accessed: 29 September 2017).

BirdLife International (2016b) *Pinguinus impennis*. The IUCN Red List of Threatened Species 2016. Available at: http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22694856A93472944.en (Accessed: 18 September 2017).

BirdLife International (2016c) *Uria lomvia.*, *The IUCN red list of threatened species*. IUCN Global Species Programme Red List Unit. Available at: http://www.iucnredlist.org/details/22694847/0 (Accessed: 30 September 2017).

BirdLife International-IUCN (2016) *BirdLife International and Handbook of the Birds of the World (2016) 2016. Pinguinus impennis. The IUCN Red List of Threatened Species. Version 2016-3, IUCN - International Union for Conservation of Nature*. Available at: http://maps.iucnredlist.org/map.html?id=22694856.

Birkhead, T. R. (1993) *Great Auk Islands : a field biologist in the Arctic*. Poyser.

Blackburn, T. M., Gaston, K. J. and Loder, N. (1999) 'Geographic gradients in body size: a clarification of Bergmann's rule.', *Diversity and Distributions*. Blackwell Science Ltd, 5(4), pp. 165–174. doi: 10.1046/j.1472-4642.1999.00046.x.

Boessenkool, S., Star, B., Scofield, R. P., Seddon, P. J. and Waters, J. M. (2010) 'Lost in translation or deliberate falsification? Genetic analyses reveal erroneous museum data for historic penguin specimens', *Proceedings of the Royal Society B: Biological Sciences*, 277(1684), pp. 1057–1064. doi: 10.1098/rspb.2009.1837.

Brace, S., Palkopoulou, E., Dalén, L., Lister, A. M., Miller, R., Otte, M., Germonpré, M., Blockley, S. P. E., Stewart, J. R. and Barnes, I. (2012) 'Serial population extinctions in a small mammal indicate Late Pleistocene ecosystem instability.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 109(50), pp. 20532–20536. doi: 10.1073/pnas.1213322109.

Brace, S., Turvey, S. T., Weksler, M., Hoogland, M. L. P. and Barnes, I. (2015) 'Unexpected evolutionary diversity in a recently extinct Caribbean mammal radiation', *Proceedings of the Royal Society of London B: Biological Sciences*, 282(1807).

Brand, S. (2016) *Is the Great Auk a Candidate for De-Extinction? — Blog of the Long Now, The Long Now Foundation Blog*. Available at: http://blog.longnow.org/02016/02/04/is-the-great-auk-a-candidate-for-de-extinction/ (Accessed: 4 September 2017).

Broad Institute (no date) *Picard Tools - By Broad Institute*. Available at: http://broadinstitute.github.io/picard/ (Accessed: 16 September 2017).

Brodkorb, P. (1960) 'Great Auk and Common Murre from a Florida Midden', *Auk*, 77(3), pp. 342–343.

de Bruyn, M., Hall, B. L., Chauke, L. F., Baroni, C., Koch, P. L. and Hoelzel, A. R. (2009) 'Rapid Response of a Marine Mammal Species to Holocene Climate and Habitat Change', *PLoS Genetics*. Cold Spring Harbor Laboratory Press, 5(7), p. e1000554. doi: 10.1371/journal.pgen.1000554.

de Bruyn, M., Hoelzel, A. R., Carvalho, G. R. and Hofreiter, M. (2011) 'Faunal histories from Holocene ancient DNA', *Trends Ecol Evol*, 26(8), pp. 405–13. doi: doi: 10.1016/j.tree.2011.03.021.

Bunce, M., Szulkin, M., Lerner, H. R. L., Barnes, I., Shapiro, B., Cooper, A. and Holdaway, R. N. (2005) 'Ancient DNA Provides New Insights into the Evolutionary History of New Zealand's Extinct Giant Eagle', *PLoS Biology*. Sinauer, 3(1), p. e9. doi: 10.1371/journal.pbio.0030009.

Burness, G. P. and Montevecchi, W. A. (1992) 'Oceanographic-related variation in the bone sizes of extinct great auks', *Polar Biology*. Springer-Verlag, 11(8), pp. 545–551. doi: 10.1007/BF00237947.

Campmas, E., Laroulandie, V., Michel, P., Amani, F., Nespoulet, R. and El Hajraoui, M. A. (2010) 'A Great Auk (*Pinguinus impennis*) in North Africa : Discovery of a bone remain in Neolithic layer of El Harhoura 2 Cave (Temara, Morocco)', in *Birds in Archaeology. Proceedings of the 6th meeting of the ICAZ bird working group in Groningen. Groningen Institute of Archaeology*, pp. 233–240.

Campos, P. F. and Gilbert, T. M. P. (2012) 'DNA Extraction from Keratin and Chitin', in Shapiro, B. and Hofreiter, M. (eds) *Ancient DNA: Methods and Protocols*. Totowa, NJ, USA: Humana Press, pp. 43–49. doi: 10.1007/978-1-61779-516-9_6.

Campos, P. F., Willerslev, E., Sher, A., Orlando, L., Axelsson, E., Tikhonov, A., Aaris-Sørensen, K., Greenwood, A. D., Kahlke, R.-D., Kosintsev, P., Krakhmalnaya, T., Kuznetsova, T., Lemey, P., MacPhee, R., Norris, C. A., Shepherd, K., Suchard, M. A., Zazula, G. D., Shapiro, B. and Gilbert, M. T. P. (2010) 'Ancient DNA analyses exclude humans as the driving force behind late Pleistocene musk ox (Ovibos moschatus) population dynamics.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 107(12), pp. 5675–5680. doi: 10.1073/pnas.0907189107.

Cano, R. J., Poinar, H. N., Pieniazek, N. J., Acra, A. and Poinar, G. O. (1993) 'Amplification and sequencing of DNA from a 120–135-million-year-old weevil', *Nature*. Nature Publishing Group, 363(6429), pp. 536–538. doi: 10.1038/363536a0.

Carøe, C., Gopalakrishnan, S., Vinner, L., Mak, S. S. T., Sinding, M.-H. S., Samaniego, J. A., Wales, N., Sicheritz-Pontén, T. and Gilbert, M. T. P. (2017) 'Single-tube library preparation for degraded DNA', *Methods in Ecology and Evolution*. doi: 10.1111/2041-210X.12871.

Carpenter, M. L., Buenrostro, J. D., Valdiosera, C., Schroeder, H., Allentoft, M. E., Sikora, M., Rasmussen, M., Gravel, S., Guillén, S., Nekhrizov, G., Leshtakov, K., Dimitrova, D., Theodossiev, N., Pettener, D., Luiselli, D., Sandoval, K., Moreno-Estrada, A., Li, Y., Wang, J., Gilbert, M. T. P., Willerslev, E., Greenleaf, W. J. and Bustamante, C. D. (2013) 'Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries.', *American Journal of Human /Genetics*. Elsevier, 93(5), pp. 852–864. doi: 10.1016/j.ajhg.2013.10.002.

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M. and Palmer, T. M. (2015) 'Accelerated modern human–induced species losses: Entering the sixth mass extinction', *Science Advances*, 1(5), p. e1400253. doi: 10.1126/sciadv.1400253.

Ceballos, G., Ehrlich, P. R. and Dirzo, R. (2017) 'Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 114(30), pp. E6089–E6096. doi: 10.1073/pnas.1704949114.

Chang, D., Knapp, M., Enk, J., Lippold, S., Kircher, M., Lister, A., Macphee, R. D. E., Widga, C., Czechowski, P., Sommer, R., Hodges, E., Stümpel, N., Barnes, I., Dalén, L., Derevianko, A., Joger, U., Poinar, H., Hofreiter, M. and Shapiro, B. (2017) 'The evolutionary and phylogeographic history of woolly mammoths: a comprehensive mitogenomic analysis', *Nature Publishing Group*, 7(44585). doi: 10.1038/srep44585.

Cooper, A. and Poinar, H. N. (2000) 'Ancient DNA: do it right or not at all.', *Science (New York, N.Y.)*, 289(5482), p. 1139. doi: 10.1126/science.289.5482.1139b.

Cooper, A., Turney, C., Hughen, K. A., Brook, B. W., McDonald, H. G. and Bradshaw, C. J. A. (2015) 'Abrupt warming events drove Late Pleistocene Holarctic megafaunal turnover', *Science*, 349(6248), pp. 602–606. doi: 10.1126/science.aac4315.

Courchamp, F., Angulo, E., Rivalan, P., Hall, R. J., Signoret, L., Bull, L. and Meinard, Y. (2006) 'Rarity Value and Species Extinction: The Anthropogenic Allee Effect', *PLoS Biology*. Public Library of Science, 4(12), p. e415. doi: 10.1371/journal.pbio.0040415.

Croxall, J. P., Butchart, S. H. M., Lascelles, B., Stattersfield, A. J., Sullivan, B., Symes, A. and Taylor, P. (2012) 'Seabird conservation status, threats and priority actions: a global assessment', *Bird Conservation International*, 22(1), pp. 1–34. doi: 10.1017/S0959270912000020.

Cuvier, G. (1813) *Essay on the theory of the earth.* Edinburgh, Printed for W. Blackwood; London, J. Murray, R. Baldwin. Available at: http://lhldigital.lindahall.org/cdm/ref/collection/earththeory/id/4519 (Accessed: 16 September 2017).

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L. and Meyer, M. (2013) 'Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 110(39), pp. 15758–63. doi: 10.1073/pnas.1314445110.

Dabney, J., Meyer, M. and Pääbo, S. (2013) 'Ancient DNA damage.', *Cold Spring Harbor perspectives in biology*. Cold Spring Harbor Laboratory Press, 5(7), p. a012567. doi: 10.1101/cshperspect.a012567.

Dalén, L., Nyström, V., Valdiosera, C., Germonpré, M., Sablin, M., Turner, E., Angerbjörn, A., Arsuaga, J. L. and Götherström, A. (2007) 'Ancient DNA reveals lack of postglacial habitat tracking in the arctic fox.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 104(16), pp. 6726–9. doi: 10.1073/pnas.0701341104.

Damgaard, P. B., Margaryan, A., Schroeder, H., Orlando, L., Willerslev, E. and Allentoft, M. E. (2015) 'Improving access to endogenous DNA in ancient bones and teeth.', *Scientific reports*. Nature Publishing Group, 5, p. 11184. doi: 10.1038/srep11184.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G. and Durbin, R. (2011) 'The variant call format and VCFtools', *Bioinformatics*. Oxford University Press, 27(15), pp. 2156–2158. doi: 10.1093/bioinformatics/btr330.

Darriba, D., Taboada, G. L., Doallo, R. and Posada, D. (2012) 'jModelTest 2: more models, new heuristics and parallel computing', *Nature Methods*, 9(8), pp. 772–772. doi: 10.1038/nmeth.2109.

Drummond, A. J. and Bouckaert, R. R. (2015) *Bayesian evolutionary analysis ; with BEAST*. Cambridge: Cambridge University Press.

Drummond, A. J., Rambaut, A., Shapiro, B. and Pybus, O. G. (2005) 'Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences', *Molecular Biology and Evolution*, 22(5), pp. 1185–1192. doi: 10.1093/molbev/msi103.

Drummond, A. J., Suchard, M. A., Xie, D. and Rambaut, A. (2012) 'Bayesian phylogenetics with BEAUti and the BEAST 1.7.', *Molecular biology and evolution*, 29(8), pp. 1969–73. doi: 10.1093/molbev/mss075.

Duchêne, S., Duchêne, D., Holmes, E. C. and Ho, S. Y. W. (2015) 'The Performance of the Date-Randomization Test in Phylogenetic Analyses of Time-Structured Virus Data', *Molecular Biology and Evolution*, 32(7), pp. 1895–1906. doi: 10.1093/molbev/msv056.

Elliott, K. H., Ricklefs, R. E., Gaston, A. J., Hatch, S. A., Speakman, J. R. and Davoren, G. K. (2013) 'High flight costs, but low dive costs, in auks support the biomechanical hypothesis for flightlessness in penguins.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 110(23), pp. 9380–4. doi: 10.1073/pnas.1304838110.

Enk, J., Rouillard, J.-M. and Poinar, H. (2013) 'Quantitative PCR as a predictor of aligned ancient DNA read counts following targeted enrichment', *BioTechniques*, 55(6), pp. 300–309. doi: 10.2144/000114114.

Espolosin Elorza, M. (2014) ‘Explotación de aves marinas en el Tardiglaciar del Golfo de Bizkaia las aves de Santa Catalina’, in *kobie .Bizkaiko Arkeologi Indusketak = Excavaciones Arqueológicas en Bizkaia*. Diputación Foral de Bizkaya, pp. 263–296. Available at: https://dialnet.unirioja.es/servlet/articulo?codigo=5218091 (Accessed: 5 September 2017).

ESRI (2016) ‘ArcGIS Desktop’. Redlands, CA, USA.

Foote, A. D., Kaschner, K., Schultze, S. E., Garilao, C., Ho, S. Y. W., Post, K., Higham, T. F. G., Stokowska, C., Van Der Es, H., Embling, C. B., Gregersen, K., Johansson, F., Willerslev, E., Thomas, M. and Gilbert, P. (2013) ‘Ancient DNA reveals that bowhead whale lineages survived Late Pleistocene climate change and habitat shifts’, *Nature Communications*, 4(1677). doi: 10.1038/ncomms2714.

Francioli, L. C., Menelaou, A., Pulit, S. L., van Dijk, F., Palamara, P. F., Elbers, C. C., Neerincx, P. B. T., Ye, K., Guryev, V., Kloosterman, W. P., Deelen, P., Abdellaoui, A., van Leeuwen, E. M., van Oven, M., Vermaat, M., Li, M., Laros, J. F. J., Karssen, L. C., Kanterakis, A., Amin, N., Hottenga, J. J., Lameijer, E.-W., Kattenberg, M., Dijkstra, M., Byelas, H., van Setten, J., van Schaik, B. D. C., Bot, J., Nijman, I. J., Renkens, I., Marschall, T., Schönhuth, A., Hehir-Kwa, J. Y., Handsaker, R. E., Polak, P., Sohail, M., Vuzman, D., Hormozdiari, F., van Enckevort, D., Mei, H., Koval, V., Moed, M. H., van der Velde, K. J., Rivadeneira, F., Estrada, K., Medina-Gomez, C., Isaacs, A., McCarroll, S. A., Beekman, M., de Craen, A. J. M., Suchiman, H. E. D., Hofman, A., Oostra, B., Uitterlinden, A. G., Willemsen, G., Study, L. C., Platteel, M., Veldink, J. H., van den Berg, L. H., Pitts, S. J., Potluri, S., Sundar, P., Cox, D. R., Sunyaev, S. R., Dunnen, J. T. den, Stoneking, M., de Knijff, P., Kayser, M., Li, Q., Li, Y., Du, Y., Chen, R., Cao, H., Li, N., Cao, S., Wang, J., Bovenberg, J. A., Pe’er, I., Slagboom, P. E., van Duijn, C. M., Boomsma, D. I., van Ommen, G.-J. B., de Bakker, P. I. W., Swertz, M. A. and Wijmenga, C. (2014) ‘Whole-genome sequence variation, population structure and demographic history of the Dutch population’, *Nature Genetics*. Nature Research, 46(8), pp. 818–825. doi: 10.1038/ng.3021.

Frankham, R. (1998) ‘Inbreeding and Extinction: Island Populations’, *Conservation Biology*. WileySociety for Conservation Biology, 12(3), pp. 665–675. doi: 10.2307/2387248.

Frankham, R. (2005) ‘Genetics and extinction’, *Biological Conservation*. Elsevier, 126(2), pp. 131–140. doi: 10.1016/J.BIOCON.2005.05.002.

Frankham, R., Bradshaw, C. J. A. and Brook, B. W. (2014) ‘Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses’, *Biological Conservation*. Elsevier, 170, pp. 56–63. doi: 10.1016/J.BIOCON.2013.12.036.

Franklin, I. R. (1980) ‘Evolutionary change in small populations’, in Soule, M. E. and Wilcox, B. A. (eds) *Conservation Biology: An Evolutionary-Ecological Perspective.* Sunderland, Mass: Sinauer Associates, pp. 135–140.

Friesen, V. L., Burd, T. M. and McCoy, K. D. (2007) 'Mechanisms of population differentiation in seabirds', *Molecular Ecology*. Blackwell Publishing Ltd, 16(9), pp. 1765–1785. doi: 10.1111/j.1365-294X.2006.03197.x.

Fu, Q., Hajdinjak, M., Moldovan, O. T., Constantin, S., Mallick, S., Skoglund, P., Patterson, N., Rohland, N., Lazaridis, I., Nickel, B., Viola, B., Prüfer, K., Meyer, M., Kelso, J., Reich, D. and Pääbo, S. (2015) 'An early modern human from Romania with a recent Neanderthal ancestor', *Nature*. Nature Research, 524(7564), pp. 216–219. doi: 10.1038/nature14558.

Fu, Q., Mittnik, A., Johnson, P. L. F., Bos, K., Lari, M., Bollongino, R., Sun, C., Giemsch, L., Schmitz, R., Burger, J., Ronchitelli, A. M., Martini, F., Cremonesi, R. G., Svoboda, J., Bauer, P., Caramelli, D., Castellano, S., Reich, D., Pääbo, S. and Krause, J. (2013) 'A Revised Timescale for Human Evolution Based on Ancient Mitochondrial Genomes', *Current Biology*, 23(7), pp. 553–559. doi: 10.1016/j.cub.2013.02.044.

Fuller, E. (1999) *The Great Auk*. Kent, England: Errol Fuller.

Fulton, T. L., Wagner, S. M., Fisher, C. and Shapiro, B. (2012) 'Nuclear DNA from the extinct Passenger Pigeon (Ectopistes migratorius) confirms a single origin of New World pigeons', *Annals of Anatomy*, 194(1), pp. 52–57. doi: 10.1016/j.aanat.2011.02.017.

Gansauge, M.-T. and Meyer, M. (2013) 'Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA', *Nature Protocols*, 8(4), pp. 737–748. doi: 10.1038/nprot.2013.038.

Ganter, B. and Gaston, A. J. (2013) *CAFF - Arctic Biodiversity Assessment 2013: Chapter 4, Birds*. Available at: https://www.caff.is/assessment-series/arctic-biodiversity-assessment/209-arctic-biodiversity-assessment-2013-chapter-4-birds.

Gaskell, J. (2000) *Who killed the great auk?* Oxford University Press.

Genome 10K Community of Scientists (2009) 'Genome 10K: A Proposal to Obtain Whole-Genome Sequence for 10 000 Vertebrate Species', *Journal of Heredity*, 100(6), pp. 659–674. doi: 10.1093/jhered/esp086.

Gilbert, M. T. P., Bandelt, H.-J., Hofreiter, M. and Barnes, I. (2005) 'Assessing ancient DNA studies', *Trends in Ecology & Evolution*, 20(10), pp. 541–544. doi: 10.1016/j.tree.2005.07.005.

Gilbert, M. T. P., Moore, W., Melchior, L., Worobey, M. and Schultz, T. (2007) 'DNA Extraction from Dry Museum Beetles without Conferring External Morphological Damage', *PLoS ONE*. Cold Spring Harbor Laboratory, 2(3), p. e272. doi: 10.1371/journal.pone.0000272.

Gilbert, M. T. P., Tomsho, L. P., Rendulic, S., Packard, M., Drautz, D. I., Sher, A., Tikhonov, A., Dalén, L., Kuznetsova, T., Kosintsev, P., Campos, P. F., Higham, T., Collins, M. J., Wilson, A. S., Shidlovskiy, F., Buigues, B., Ericson, P. G. P., Germonpré, M., Götherström, A., Iacumin, P., Nikolaev, V., Nowak-Kemp, M., Willerslev, E., Knight, J. R., Irzyk, G. P., Perbost, C. S., Fredrikson, K. M., Harkins, T. T., Sheridan, S., Miller, W. and Schuster, S. C. (2007) 'Whole-Genome Shotgun Sequencing of Mitochondria from Ancient Hair Shafts', *Science*, 317(5846). doi: 10.1126/science.1146971.

Gouy, M., Guindon, S. and Gascuel, O. (2010) 'SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building', *Molecular Biology and Evolution*. Oxford University Press, 27(2), pp. 221–224. doi: 10.1093/molbev/msp259.

GraphPad (no date) *GraphPad QuickCalcs*. Available at: https://www.graphpad.com/quickcalcs/ttest1.cfm?Format=SD (Accessed: 1 September 2017).

Grealy, A., Phillips, M., Miller, G., Gilbert, M. T. P., Rouillard, J.-M., Lambert, D., Bunce, M. and Haile, J. (2017a) 'Eggshell palaeogenomics: Palaeognath evolutionary history revealed through ancient nuclear and mitochondrial DNA from Madagascan elephant bird (Aepyornis sp.) eggshell', *Molecular Phylogenetics and Evolution*, 109, pp. 151–163. doi: 10.1016/j.ympev.2017.01.005.

Grealy, A., Rawlence, N. J. and Bunce, M. (2017b) 'Time to spread your wings: A review of the avian ancient DNA field', *Genes*. Multidisciplinary Digital Publishing Institute, p. 184. doi: 10.3390/genes8070184.

Green, E. J. and Speller, C. F. (2017) 'Novel Substrates as Sources of Ancient DNA: Prospects and Hurdles', *Genes*, 8(7), p. 180. doi: 10.3390/genes8070180.

Green, R. E., Krause, J., Ptak, S. E., Briggs, A. W., Ronan, M. T., Simons, J. F., Du, L., Egholm, M., Rothberg, J. M., Paunovic, M. and Pääbo, S. (2006) 'Analysis of one million base pairs of Neanderthal DNA', *Nature*. Nature Publishing Group, 444(7117), pp. 330–336. doi: 10.1038/nature05336.

Greenwood, A. D., Capelli, C., Possnert, G. and Paabo, S. (1999) 'Nuclear DNA sequences from late Pleistocene megafauna', *Molecular Biology and Evolution*. Oxford University Press, 16(11), pp. 1466–1473. doi: 10.1093/oxfordjournals.molbev.a026058.

Grieve, S. (1885) *The Great Auk, or Garefowl, its history, archaeology and remains*. Digitally. Cambridge, United Kingdom: Cambridge University Pess.

Groot, M. (2005) 'The Great Auk (*Pinguinus impennis*) in the Netherlands during the Roman Period', *International Journal of Osteoarchaeology*, 15(1), pp. 15–22. doi: 10.1002/oa.743.

Guindon, S. and Gascuel, O. (2003) 'A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood', *Systematic Biology*. Oxford University Press, 52(5), pp. 696–704. doi: 10.1080/10635150390235520.

Gutiérrez-García, T. A., Vázquez-Domínguez, E., Arroyo-Cabrales, J., Kuch, M., Enk, J., King, C. and Poinar, H. N. (2014) 'Ancient DNA and the tropics: a rodent's tale', *Biology Letters*, 10(6). doi: 10.1098/rsbl.2014.0224.

Hadly, E. A., Ramakrishnan, U., Chan, Y. L., van Tuinen, M., O'Keefe, K., Spaeth, P. A. and Conroy, C. J. (2004) 'Genetic response to climatic change: insights from ancient DNA and phylochronology.', *PLoS biology*. Public Library of Science, 2(10), p. e290. doi: 10.1371/journal.pbio.0020290.

Hagelberg, E., Hofreiter, M. and Keyser, C. (2014) 'Ancient DNA: the first three decades', *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370(1660).

Hagelberg, E., Sykes, B. and Hedges, R. (1989) 'Ancient bone DNA amplified', *Nature*, 342(6249), pp. 485–485. doi: 10.1038/342485a0.

Hall, T. A. (1999) 'Bioedit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT'. NUCLEIC ACIDS SYMPOSIUM SERIES, pp. 95–98.

Halliday, T. (1978) *Vanishing birds: their natural history and conservation*. Holt, Rinehart and Winston.

Halliday, T. R. (1980) 'The extinction of the passenger pigeon ectopistes migratorius and its relevance to contemporary conservation', *Biological Conservation*. Elsevier, 17(2), pp. 157–162. doi: 10.1016/0006-3207(80)90046-4.

Hansen, H. B., Damgaard, P. B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E. and Allentoft, M. E. (2017) 'Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum', *PLOS ONE*. Public Library of Science, 12(1), p. e0170940. doi: 10.1371/journal.pone.0170940.

Hasegawa, M., Kishino, H. and Yano, T. (1985) 'Dating of the human-ape splitting by a molecular clock of mitochondrial DNA.', *Journal of molecular evolution*, 22(2), pp. 160–74.

Hay, O. P. (1902) 'On the Finding of the Bones of the Great Auk (*Plautus impennis*) in Florida', *The Auk*, 19(3), pp. 255–258. doi: 10.2307/4069020.

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O. A. and Wilson, A. C. (1984) 'DNA sequences from the quagga, an extinct member of the horse family', *Nature*, 312(5991), pp. 282–284. doi: 10.1038/312282a0.

Hobson, K. A. and Montevecchi, W. A. (1991) 'Stable isotopic determinations of trophic relationships of great auks', *Oecologia*. Springer-Verlag, 87(4), pp. 528–531. doi: 10.1007/BF00320416.

Hofreiter, M., Paijmans, J. L. A., Goodchild, H., Speller, C. F., Barlow, A., Fortes, G. G., Thomas, J. A., Ludwig, A. and Collins, M. J. (2015) 'The future of ancient DNA: Technical advances and conceptual shifts', *BioEssays*, 37(3), pp. 284–293. doi: 10.1002/bies.201400160.

Hofreiter, M., Rabeder, G., Jaenicke-Despré, V., Withalm, G., Nagel, D., Paunovic, M., Jambre, G. and Pä ä bo, S. (2004) 'Evidence for Reproductive Isolation between Cave Bear Populations', *Current Biology*, 14, pp. 40–43. doi: 10.1016/j.cub.2003.12.035.

Hofreiter, M., Serre, D., Poinar, H. N., Kuch, M. and Pääbo, S. (2001) 'Ancient DNA', *Nature Reviews Genetics*, 2(5), pp. 353–359. doi: 10.1038/35072071.

Holdaway, R. N., Allentoft, M. E., Jacomb, C., Oskam, C. L., Beavan, N. R. and Bunce, M. (2014) 'An extremely low-density human population exterminated New Zealand moa', *Nature Communications*, 5(5436). doi: 10.1038/ncomms6436.

Horai, S., Hayasaka, K., Murayama, K. and Wate, N. (1989) 'DNA amplification from ancient human skeletal remains and their sequence analysis', *Proceedings of the Japan Academy, Series B*, 65(10), pp. 229–233. doi: 10.2183/pjab.65.229.

Houston, A. I., Wood, J. and Wilkinson, M. (2010) 'How did the Great Auk raise its young?', *Journal of Evolutionary Biology*. Blackwell Publishing Ltd, 23(9), pp. 1899–1906. doi: 10.1111/j.1420-9101.2010.02047.x.

Hufthammer, A. K. (1982) *Geirfuglens utbredelse og morfologiske variasjon i Skandinavia*. Universitetet i Bergen.

Hume, J. P. (2006) 'The history of the Dodo *Raphus cucullatus* and the penguin of Mauritius', *Historical Biology*, 18(2), pp. 65–89. doi: 10.1080/08912960600639400.

Hung, C.-M., Shaner, P.-J. L., Zink, R. M., Liu, W.-C., Chu, T.-C., Huang, W.-S. and Li, S.-H. (2014) 'Drastic population fluctuations explain the rapid extinction of the passenger pigeon.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 111(29), pp. 10636–41. doi: 10.1073/pnas.1401526111.

Huynen, L., Millar, C. D., Scofield, R. P., Lambert, D. and Nature (2003) 'Nuclear DNA sequences detect species limits in ancient moa', *Nature*, 425(6954), pp. 175–178. doi: 10.1038/nature01838.

IBM Corp (2016) 'IBM SPSS Statistics for Windows'. Armonk, NY: IBM Crop.

Jain, S., Rai, N., Kumar, G., Pruthi, P. A., Thangaraj, K., Bajpai, S. and Pruthi, V. (2017) 'Ancient DNA Reveals Late Pleistocene Existence of Ostriches in Indian Sub-Continent', *PLOS ONE*. Geological survey of Namibia, 12(3), p. e0164823. doi: 10.1371/journal.pone.0164823.

Janoo, A. (2005) ‘Discovery of isolated dodo bones [*Raphus cucullatus* (L.), Aves, Columbiformes] from Mauritius cave shelters highlights human predation, with a comment on the status of the family Raphidae Wetmore, 1930’, *Annales de Paléontologie*. Elsevier Masson, 91(2), pp. 167–180. doi: 10.1016/J.ANNPAL.2004.12.002.

Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., Ho, S. Y. W., Faircloth, B. C., Nabholz, B., Howard, J. T., Suh, A., Weber, C. C., da Fonseca, R. R., Li, J., Zhang, F., Li, H., Zhou, L., Narula, N., Liu, L., Ganapathy, G., Boussau, B., Bayzid, M. S., Zavidovych, V., Subramanian, S., Gabaldón, T., Capella-Gutiérrez, S., Huerta-Cepas, J., Rekepalli, B., Munch, K., Schierup, M., Lindow, B., Warren, W. C., Ray, D., Green, R. E., Bruford, M. W., Zhan, X., Dixon, A., Li, S., Li, N., Huang, Y., Derryberry, E. P., Bertelsen, M. F., Sheldon, F. H., Brumfield, R. T., Mello, C. V, Lovell, P. V, Wirthlin, M., Schneider, M. P. C., Prosdocimi, F., Samaniego, J. A., Vargas Velazquez, A. M., Alfaro-Núñez, A., Campos, P. F., Petersen, B., Sicheritz-Ponten, T., Pas, A., Bailey, T., Scofield, P., Bunce, M., Lambert, D. M., Zhou, Q., Perelman, P., Driskell, A. C., Shapiro, B., Xiong, Z., Zeng, Y., Liu, S., Li, Z., Liu, B., Wu, K., Xiao, J., Yinqi, X., Zheng, Q., Zhang, Y., Yang, H., Wang, J., Smeds, L., Rheindt, F. E., Braun, M., Fjeldsa, J., Orlando, L., Barker, F. K., Jønsson, K. A., Johnson, W., Koepfli, K.-P., O’Brien, S., Haussler, D., Ryder, O. A., Rahbek, C., Willerslev, E., Graves, G. R., Glenn, T. C., McCormack, J., Burt, D., Ellegren, H., Alström, P., Edwards, S. V, Stamatakis, A., Mindell, D. P., Cracraft, J., Braun, E. L., Warnow, T., Jun, W., Gilbert, M. T. P. and Zhang, G. (2014) ‘Whole-genome analyses resolve early branches in the tree of life of modern birds.’, *Science (New York, N.Y.)*. American Association for the Advancement of Science, 346(6215), pp. 1320–31. doi: 10.1126/science.1253451.

Jetz, W., Wilcove, D. S., Dobson, A. P., Leemans, R. and Monserud, R. (2007) ‘Projected Impacts of Climate and Land-Use Change on the Global Diversity of Birds’, *PLoS Biology*. Edited by G. M. Mace. U.S. Geological Survey, 5(6), p. e157. doi: 10.1371/journal.pbio.0050157.

Johnson, K. P., Clayton, D. H., Dumbacher, J. P. and Fleischer, R. C. (2010) ‘The flight of the Passenger Pigeon: Phylogenetics and biogeographic history of an extinct species’, *Molecular Phylogenetics and Evolution*. Academic Press, 57(1), pp. 455–458. doi: 10.1016/J.YMPEV.2010.05.010.

Jonsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. and Orlando, L. (2013) ‘mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters’, *Bioinformatics*, 29(13), pp. 1682–1684. doi: 10.1093/bioinformatics/btt193.

Kass, R. E. and Raftery, A. E. (1995) ‘Bayes Factors’, *Journal of the American Statistical Association*. Taylor & Francis, Ltd.American Statistical Association, 90(430), p. 773. doi: 10.2307/2291091.

Keane, T. M., Creevey, C. J., Pentony, M. M., Naughton, T. J. and Mclnerney, J. O. (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(1), p. 29. doi: 10.1186/1471-2148-6-29.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. and Drummond, A. (2012) 'Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.', *Bioinformatics*. Oxford University Press, 28(12), pp. 1647–1649. doi: 10.1093/bioinformatics/bts199.

Kehlmaier, C., Barlow, A., Hastings, A. K., Vamberger, M., Paijmans, J. L. A., Steadman, D. W., Albury, N. A., Franz, R., Hofreiter, M. and Fritz, U. (2017) 'Tropical ancient DNA reveals relationships of the extinct Bahamian giant tortoise *Chelonoidis alburyorum*', *Proceedings of the Royal Society of London B: Biological Sciences*, 284(1846). doi: 10.1098/rspb.2016.2235.

Kirkham, I. R. and Montevecchi, W. A. (1982) 'The breeding birds of Funk Island, Newfoundland: an historical perspective', *American Birds*, 36(2), pp. 111–118.

Knapp, M., Clarke, A. C., Horsburgh, K. A. and Matisoo-Smith, E. A. (2012) 'Setting the stage – Building and working in an ancient DNA laboratory', *Annals of Anatomy - Anatomischer Anzeiger*, 194(1), pp. 3–6. doi: 10.1016/j.aanat.2011.03.008.

Knapp, M. and Hofreiter, M. (2010) 'Next Generation Sequencing of Ancient DNA: Requirements, Strategies and Perspectives', *Genes*, 1(2), pp. 227–243. doi: 10.3390/genes1020227.

Ksepka, D. T., Ware, J. L. and Lamm, K. S. (2014) 'Flying rocks and flying clocks: disparity in fossil and molecular dates for birds.', *Proceedings. Biological sciences*. The Royal Society, 281(1788). doi: 10.1098/rspb.2014.0677.

Kuhlwilm, M., Gronau, I., Hubisz, M. J., de Filippo, C., Prado-Martinez, J., Kircher, M., Fu, Q., Burbano, H. A., Lalueza-Fox, C., de la Rasilla, M., Rosas, A., Rudan, P., Brajkovic, D., Kucan, Ž., Gušic, I., Marques-Bonet, T., Andrés, A. M., Viola, B., Pääbo, S., Meyer, M., Siepel, A. and Castellano, S. (2016) 'Ancient gene flow from early modern humans into Eastern Neanderthals', *Nature*. Nature Research, 530(7591), pp. 429–433. doi: 10.1038/nature16544.

Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets', *Molecular Biology and Evolution*, 33(7), pp. 1870–1874. doi: 10.1093/molbev/msw054.

Lacy, R. C. and Pollak, J. P. (2014) 'Vortex: A Stochastic Simulation of the Extinction Process. Version 10.0', *Chicago Zoological Society*.

Lagerholm, V. K., Sandoval-Castellanos, E., Vaniscotte, A., Potapova, O. R., Tomek, T., Bochenski, Z. M., Shepherd, P., Barton, N., Van Dyck, Marie-C., Miller, R., Höglund, Jacob, Yoccoz, N. G., Dalén, L. and Stewart, J. R. (2017) 'Range shifts or extinction? Ancient DNA and distribution modelling reveal past and future responses to climate warming in cold-adapted birds', *Global Change Biology*, 23(4), pp. 1425–1435. doi: 10.1111/gcb.13522.

Lambert, D. M., Ritchie, P. A., Millar, C. D., Holland, B., Drummond, A. J. and Baroni, C. (2002) 'Rates of Evolution in Ancient DNA from Adélie Penguins', *Science*, 295(5563), pp. 2270–2273. doi: 10.1126/science.1068105.

Laroulandie, V., Elorza, M. and Berganza, E. (2016) 'Les oiseaux marins du Magdalénien supérieur de Santa Catalina (Lekeitio, Biscaye, Espagne) : approches taphonomique et archéozoologique', in Dupont, C. and Marchand, G. (eds) *Archéologie des chasseurs-cueilleurs maritimes. De la fonction des habitats à l'organisation de l'espace littoral, Actes de la séance de la Société préhistorique française de Renne, Seapeople 2014, 10-11 avril 2014*. Paris: Société préhistorique française (Séances SPF 6), pp. 37–57.

Larson, G., Albarella, U., Dobney, K., Rowley-Conwy, P., Schibler, J., Tresset, A., Vigne, J.-D., Edwards, C. J., Schlumbaum, A., Dinu, A., Balaçsescu, A., Dolman, G., Tagliacozzo, A., Manaseryan, N., Miracle, P., Van Wijngaarden-Bakker, L., Masseti, M., Bradley, D. G. and Cooper, A. (2007) 'Ancient DNA, pig domestication, and the spread of the Neolithic into Europe.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 104(39), pp. 15276–15281. doi: 10.1073/pnas.0703411104.

Leeton, P., Christidis, L. and Westerman, M. (1993) 'Feathers from museum bird skins-a good source of DNA for phylogenetic studies'', *The Condor*, 95(2), pp. 465–466. doi: 10.2307/1369369.

Leigh, J. W. and Bryant, D. (2015) 'PopART: Full-feature software for haplotype network construction', *Methods in Ecology and Evolution*. Edited by S. Nakagawa, 6(9), pp. 1110–1116. doi: 10.1111/2041-210X.12410.

Leonard, J. (2008) 'Ancient DNA applications for wildlife conservation', *Molecular Ecology*, 17, pp. 4186–4196. doi: 10.1111/j.1365-294X.2008.03891.x.

Li, H. (2013) 'Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM', *Genomics*, 1303.

Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform', *Bioinformatics*, 25(14), pp. 1754–1760. doi: 10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and 1000 Genome Project Data Processing Subgroup (2009) 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*, 25(16), pp. 2078–2079. doi: 10.1093/bioinformatics/btp352.

Lindgreen, S. (2012) 'AdapterRemoval: easy cleaning of next-generation sequencing reads', *BMC Res Notes*, 5. doi: 10.1186/1756-0500-5-337.

Lindgreen, S. (2012) 'AdapterRemoval: Easy Cleaning of Next Generation Sequencing Reads', *BMC Research Notes*, 5(1), p. 337. doi: 10.1186/1756-0500-5-337.

Livezey, B. C. (1988) 'Morphometrics of flightlessness in the Alcidae', *The Auk*, 105, pp. 681–698.

Lorenzen, E. D., Nogués-Bravo, D., Orlando, L., Weinstock, J., Binladen, J., Marske, K. A., Ugan, A., Borregaard, M. K., Gilbert, M. T. P., Nielsen, R., Ho, S. Y. W., Goebel, T., Graf, K. E., Byers, D., Stenderup, J. T., Rasmussen, M., Campos, P. F., Leonard, J. A., Koepfli, K.-P., Froese, D., Zazula, G., Stafford, T. W., Aaris-Sørensen, K., Batra, P., Haywood, A. M., Singarayer, J. S., Valdes, P. J., Boeskorov, G., Burns, J. A., Davydov, S. P., Haile, J., Jenkins, D. L., Kosintsev, P., Kuznetsova, T., Lai, X., Martin, L. D., McDonald, H. G., Mol, D., Meldgaard, M., Munch, K., Stephan, E., Sablin, M., Sommer, R. S., Sipko, T., Scott, E., Suchard, M. A., Tikhonov, A., Willerslev, R., Wayne, R. K., Cooper, A., Hofreiter, M., Sher, A., Shapiro, B., Rahbek, C. and Willerslev, E. (2011) 'Species-specific responses of Late Quaternary megafauna to climate and humans', *Nature*, 479(7373), pp. 359–364. doi: 10.1038/nature10574.

Lucas, F. A. (1890) *The expedition to the Funk Island, with observations upon the history and anatomy of the great auk*. Washington : G.P.O.,. doi: 10.5962/bhl.title.44449.

Lugli, G. A., Milani, C., Mancabelli, L., Turroni, F., Ferrario, C., Duranti, S., van Sinderen, D. and Ventura, M. (2017) 'Ancient bacteria of the Ötzi's microbiome: a genomic tale from the Copper Age.', *Microbiome*. BioMed Central, 5(1), p. 5. doi: 10.1186/s40168-016-0221-y.

MacDonald, G. M., Bennett, K. D., Jackson, S. T., Parducci, L., Smith, F. A., Smol, J. P. and Willis, K. J. (2008) 'Impacts of climate change on species, populations and communities: palaeobiogeographical insights and frontiers', *Progress in Physical Geography*. SAGE Publications, 32(2), pp. 139–172. doi: 10.1177/0309133308094081.

Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akçakaya, H. R., Leader-Williams, N., Milner-Gulland, E. J. and Stuart, S. N. (2008) 'Quantification of Extinction Risk: IUCN's System for Classifying Threatened Species', *Conservation Biology*. Blackwell Publishing Inc, 22(6), pp. 1424–1442. doi: 10.1111/j.1523-1739.2008.01044.x.

Machado, N. and Loyola, R. D. (2013) 'A Comprehensive Quantitative Assessment of Bird Extinction Risk in Brazil', *PLoS ONE*. Edited by D. L. Roberts. Public Library of Science, 8(8), p. e72283. doi: 10.1371/journal.pone.0072283.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. and DePristo, M. A. (2010) 'The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data', *Genome Research*, 20(9), pp. 1297–1303. doi: 10.1101/gr.107524.110.

McKinney, M. L. (1997) 'Extinction vulnerability and selectivity: combining ecological and paleontological views', *Annual Review of Ecology and Systematics*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, 28(1), pp. 495–516. doi: 10.1146/annurev.ecolsys.28.1.495.

Meiri, S. and Dayan, T. (2003) 'On the validity of Bergmann's rule', 30(3), pp. 331–351. doi: 10.1046/j.1365-2699.2003.00837.x.

Meiri, S., Yom-Tov, Y. and Geffen, E. (2007) 'What determines conformity to Bergmann's rule?', *Global Ecology and Biogeography*. Blackwell Publishing Ltd, 16(6), pp. 788–794. doi: 10.1111/j.1466-8238.2007.00330.x.

Meldgaard, M. (1988) 'The Great Auk, *Pinguinus impennis* (L.) in Greenland', *Historical Biology*. Taylor & Francis Group , 1(2), pp. 145–178. doi: 10.1080/08912968809386472.

Merkel, F. and Barry, T. (2008) *Seabird Harvest in the Arctic*. Conservation of Arctic Flora and Fauna (CAFF). Available at: https://oaarchive.arctic-council.org/handle/11374/190 (Accessed: 16 September 2017).

Metcalf, J. L., Turney, C., Barnett, R., Martin, F., Bray, S. C., Vilstrup, J. T., Orlando, L., Salas-Gismondi, R., Loponte, D., Medina, M., De Nigris, M., Civalero, T., Fernández, P. M., Gasco, A., Duran, V., Seymour, K. L., Otaola, C., Gil, A., Paunero, R., Prevosti, F. J., Bradshaw, C. J. A., Wheeler, J. C., Borrero, L., Austin, J. J. and Cooper, A. (2016) 'Synergistic roles of climate warming and human occupation in Patagonian megafaunal extinctions during the Last Deglaciation', *Science Advances*, 2(6).

Meyer, M. and Kircher, M. (2010) 'Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing', *Cold Spring Harbor Protocols*, 2010(6), p. pdb.prot5448-prot5448. doi: 10.1101/pdb.prot5448.

Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J. G., Jay, F., Prüfer, K., de Filippo, C., Sudmant, P. H., Alkan, C., Fu, Q., Do, R., Rohland, N., Tandon, A., Siebauer, M., Green, R. E., Bryc, K., Briggs, A. W., Stenzel, U., Dabney, J., Shendure, J., Kitzman, J., Hammer, M. F., Shunkov, M. V., Derevianko, A. P., Patterson, N., Andrés, A. M., Eichler, E. E., Slatkin, M., Reich, D., Kelso, J. and Pääbo, S. (2012) 'A High-Coverage Genome Sequence from an Archaic Denisovan Individual', *Science*, 338(6104), pp. 222–226. doi: 10.1126/science.1224344.

Miller, W., Drautz, D. I., Ratan, A., Pusey, B., Qi, J., Lesk, A. M., Tomsho, L. P., Packard, M. D., Zhao, F., Sher, A., Tikhonov, A., Raney, B., Patterson, N., Lindblad-Toh, K., Lander, E. S., Knight, J. R., Irzyk, G. P., Fredrikson, K. M., Harkins, T. T., Sheridan, S., Pringle, T. and Schuster, S. C. (2008) 'Sequencing the nuclear genome of the extinct woolly mammoth', *Nature*. Nature Publishing Group, 456(7220), pp. 387–390. doi: 10.1038/nature07446.

Milne, I., Stephen, G., Bayer, M., Cock, P. J. A., Pritchard, L., Cardle, L., Shaw, P. D. and Marshall, D. (2013) 'Using Tablet for visual exploration of second-generation sequencing data', *Briefings in Bioinformatics*. Oxford University Press, 14(2), pp. 193–202. doi: 10.1093/bib/bbs012.

Minin, V. N., Bloomquist, E. W. and Suchard, M. A. (2008) 'Smooth Skyride through a Rough Skyline: Bayesian Coalescent-Based Inference of Population Dynamics', *Molecular Biology and Evolution*, 25(7), pp. 1459–1471. doi: 10.1093/molbev/msn090.

Mitchell, K. J., Llamas, B., Soubrier, J., Rawlence, N. J., Worthy, T. H., Wood, J., Lee, M. S. Y. and Cooper, A. (2014) 'Ancient DNA reveals elephant birds and kiwi are sister taxa and clarifies ratite bird evolution', *Science*, 344(6186), pp. 898–900. doi: 10.1126/science.1251981.

Moen, S. M. (1991) 'Morphologic and genetic variation among breeding colonies of the Atlantic puffin (*Fratercula arctica*)', *The Auk*, 108(4), pp. 755–763.

Mohandesan, E., Speller, C. F., Peters, J., Uerpmann, H.-P., Uerpmann, M., De Cupere, B., Hofreiter, M. and Burger, P. A. (2017) 'Combined hybridization capture and shotgun sequencing for ancient DNA analysis of extinct wild and domestic dromedary camel', *Molecular Ecology Resources*, 17(2), pp. 300–313. doi: 10.1111/1755-0998.12551.

Montevecchi, W. A., Chaffey, H. and Burke, C. (2007) 'Hunting for security: Changes in the exploitation of Marine birds in Newfoundland and Labrador.', in Parrish, C. C., Turner, N. J., and Solberg, S. M. (eds) *Resetting the Kitchen Table*. Nova Science Publishers, Inc, pp. 99–116.

Montevecchi, W. A. and Kirk, D. A. (1996) 'Great Auk (*Pinguinus impennis*)', *The Birds of North America Online*. doi: 10.2173/bna.260.

Morris-Pocock, J. A., Taylor, S. A., Birt, T. P., Damus, M., Piatt, J. F., Warheit, K. I. and Friesen, V. L. (2008) 'Population genetic structure in Atlantic and Pacific Ocean common murres (*Uria aalge*): natural replicate tests of post-Pleistocene evolution', *Molecular Ecology*. Blackwell Publishing Ltd, 17(22), pp. 4859–4873. doi: 10.1111/j.1365-294X.2008.03977.x.

Moum, T. and Árnason, E. (2001) 'Genetic diversity and population history of two related seabird species based on mitochondrial DNA control region sequences', *Molecular Ecology*. Blackwell Science Ltd, 10(10), pp. 2463–2478. doi: 10.1046/j.0962-1083.2001.01375.x.

Moum, T., Árnason, U. and Árnason, E. (2002) 'Mitochondrial DNA sequence evolution and phylogeny of the Atlantic Alcidae, including the extinct great auk (*Pinguinus impennis*).', *Molecular biology and evolution*, 19(9), pp. 1434–9.

Murray, G. G. R., Soares, A. E. R., Novak, B. J., Schaefer, N. K., Cahill, J. A., Baker, A. J., Demboski, J. R., Doll, A., Da Fonseca, R. R., Fulton, T. L., Gilbert, M. T. P., Heintzman, P. D., Letts, B., McIntosh, G., O'Connell, B. L., Peck, M., Pipes, M.-L., Rice, E. S., Santos, K. M., Sohrweide, A. G., Vohr, S. H., Corbett-Detig, R. B., Green, R. E. and Shapiro, B. (2017) 'Natural selection shaped the rise and fall of passenger pigeon genomic diversity.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 358(6365), pp. 951–954. doi: 10.1126/science.aao0960.

MYcroarray (2016) *MYbaits- Frequently Asked Questions*. Available at: http://www.mycroarray.com/mybaits/mybaits-faq.html (Accessed: 26 September 2017).

MYcroarray MYbaits (2014) *MYcroarray MYbaits Sequence Enrichment for Targeted Sequencing User Manual*. Available at: http://www.mycroarray.com/pdf/MYbaits-manual-v2.pdf (Accessed: 17 May 2017).

MYcroarray and MYbaits (2016) *Manual for MYbaits Target Enrichment Kit v3*. Available at: http://www.mycroarray.com/mybaits/manuals.html (Accessed: 15 September 2017).

Nabholz, B., Lanfear, R. and Fuchs, J. (2016) 'Body mass-corrected molecular rate for bird mitochondrial DNA', *Molecular Ecology*, 25(18), pp. 4438–4449. doi: 10.1111/mec.13780.

Newton, A. (1861) 'XLII.-Abstract of Mr. J. Wolley's researches in Iceland respecting the Gare-fowl or Great Auk (*Alca impennis*, Linn.)', *Ibis*. Blackwell Publishing Ltd, 3(4), pp. 374–399. doi: 10.1111/j.1474-919X.1861.tb08857.x.

O'Corry-Crowe, G. (2008) 'Climate change and the molecular ecology of Arctic marine mammals.', *Ecological applications*, 18(2 Suppl), pp. S56-76.

Olson, S. L. (1977) 'A great auk, *Pinguinis* [sic], from the Pliocene of North Carolina (Aves: Alcidae)', *Proceedings of the Biological Society of Washington*, 90(3), pp. 690–697.

Olson, S. L. (1989) 'Extinction on islands: man as a catastrophe', in Western, D. and Pearl, M. (eds) *Conservation Biology for the Next Century*. Oxford University Press, pp. 50–53.

Olson, S. L., Swift, C. C. and Mokhiber, C. (1979) 'An Attempt to Determine the Prey of the Great Auk (*Pinguinus impennis*)', *The Auk*. American Ornithological Society, pp. 790–792. doi: 10.2307/4085666.

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P. L. F., Fumagalli, M., Vilstrup, J. T., Raghavan, M., Korneliussen, T., Malaspinas, A.-S., Vogt, J., Szklarczyk, D., Kelstrup, C. D., Vinther, J., Dolocan, A., Stenderup, J., Velazquez, A. M. V, Cahill, J., Rasmussen, M., Wang, X., Min, J., Zazula, G. D., Seguin-Orlando, A., Mortensen, C., Magnussen, K., Thompson, J. F., Weinstock, J., Gregersen, K., Roed, K. H., Eisenmann, V., Rubin, C. J., Miller, D. C., Antczak, D. F., Bertelsen, M. F., Brunak, S., Al-Rasheid, K. A. S., Ryder, O., Andersson, L., Mundy, J., Krogh, A., Gilbert, M. T. P., Kjaer, K., Sicheritz-Ponten, T., Jensen, L. J., Olsen, J. V, Hofreiter, M., Nielsen, R., Shapiro, B., Wang, J. and Willerslev, E. (2013) 'Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse', *Nature*, 499(7456), pp. 74–78. doi: 10.1038/nature12323.

Orlando, L., Metcalf, J. L., Alberdi, M. T., Telles-Antunes, M., Bonjean, D., Otte, M., Martin, F., Eisenmann, V., Mashkour, M., Morello, F., Prado, J. L., Salas-Gismondi, R., Shockey, B. J., Wrinn, P. J., Vasil'ev, S. K., Ovodov, N. D., Cherry, M. I., Hopwood, B., Male, D., Austin, J. J., Hänni, C. and Cooper, A. (2009) 'Revising the recent evolutionary history of equids using ancient DNA.', *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 106(51), pp. 21754–9. doi: 10.1073/pnas.0903672106.

Oskam, C. L., Haile, J., McLay, E., Rigby, P., Allentoft, M. E., Olsen, M. E., Bengtsson, C., Miller, G. H., Schwenninger, J.-L., Jacomb, C., Walter, R., Baynes, A., Dortch, J., Parker-Pearson, M., Gilbert, M. T. P., Holdaway, R. N., Willerslev, E. and Bunce, M. (2010) 'Fossil avian eggshell preserves ancient DNA.', *Proceedings. Biological sciences*. The Royal Society, 277(1690), pp. 1991–2000. doi: 10.1098/rspb.2009.2019.

Ottoni, C., Neer, W. Van, De Cupere, B., Daligault, J., Guimaraes, S., Peters, J., Spassov, N., Prendergast, M. E., Boivin, N., Morales-Muñiz, A., Bălăşescu, A., Becker, C., Benecke, N., Crowther, A., Llorente, L., Manaseryan, N., Monchot, H., Onar, V., Osypińska, M., Putelat, O., Morales, E. M. Q., Studer, J., Wierer, U., Decorte, R., Grange, T. and Geigl, E.-M. (2017) 'The palaeogenetics of cat dispersal in the ancient world', *Nature Ecology & Evolution*, 1(139). doi: 10.1038/s41559-017-0139.

Owens, I. P. F. and Bennett, P. M. (2000) 'Ecological basis of extinction risk in birds: Habitat loss versus human persecution and introduced predators', *Proceedings of the National Academy of Sciences*, 97(22), pp. 12144–12148. doi: 10.1073/pnas.200223397.

Pääbo, S. (1985) 'Preservation of DNA in ancient Egyptian mummies', *Journal of Archaeological Science*. Academic Press, 12(6), pp. 411–417. doi: 10.1016/0305-4403(85)90002-0.

Pääbo, S. (1989) 'Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification.', *Proceedings of the National Academy of Sciences*, 86(6), pp. 1939–43.

Pääbo, S., Gifford, J. A. and Wilson, A. C. (1988) 'Mitochondrial DNA sequences from a 7000-year old brain', *Nucleic Acids Research*. Oxford University Press, 16(20), pp. 9775–9787. doi: 10.1093/nar/16.20.9775.

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L. and Hofreiter, M. (2004) 'Genetic Analyses from Ancient DNA', *Annual Review of Genetics*, 38(1), pp. 645–679. doi: 10.1146/annurev.genet.37.110801.143214.

Paleczny, M., Hammill, E., Karpouzi, V., Pauly, D., Veitch, D. and Puerta, N. (2015) 'Population Trend of the World's Monitored Seabirds, 1950-2010', *PLOS ONE*. Edited by M. Krkosek. Nelson College, 10(6), p. e0129342. doi: 10.1371/journal.pone.0129342.

Palumbi, S. R. and Baker, C. S. (1994) 'Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales.', *Molecular Biology and Evolution*. Oxford University Press, 11(3), pp. 426–435. doi: 10.1093/oxfordjournals.molbev.a040115.

Pan, H., Yu, H., Ravi, V., Li, C., Lee, A. P., Lian, M. M., Tay, B.-H., Brenner, S., Wang, J., Yang, H., Zhang, G. and Venkatesh, B. (2016) 'The genome of the largest bony fish, ocean sunfish *(Mola mola*), provides insights into its fast growth rate', *GigaScience*. BioMed Central, 5(1), p. 36. doi: 10.1186/s13742-016-0144-3.

Payne, R. B. and Sorenson, M. D. (2003) ‘Museum Collections as Sources of Genetic Data’, *Bonner zoologische Beiträge*, 51, pp. 97–104.

Pearce, R. L., Wood, J. J., Artukhin, Y., Birt, T. P., Damus, M. and Friesen, V. L. (2009) ‘Mitochondrial DNA suggests high gene flow in ancient murrelets’, *The C*, 104(1), pp. 84–91. doi: 10.1650/0010-5422(2002)104[0084:MDSHGF]2.0.CO;2.

Pereira, S. L. and Baker, A. J. (2008) ‘DNA evidence for a Paleocene origin of the Alcidae (Aves: Charadriiformes) in the Pacific and multiple dispersals across northern oceans’, *Molecular Phylogenetics and Evolution*, 46(2), pp. 430–445. doi: 10.1016/j.ympev.2007.11.020.

Perry, G. L. W., Wheeler, A. B., Wood, J. R. and Wilmshurst, J. M. (2014) ‘A high-precision chronology for the rapid extinction of New Zealand moa (Aves, Dinornithiformes)’, *Quaternary Science Reviews*. Pergamon, 105, pp. 126–135. doi: 10.1016/J.QUASCIREV.2014.09.025.

Piatt, J. F. and Nettleship, D. N. (1985) ‘Diving Depths of Four Alcids’, *The Auk*. American Ornithological Society, 102(2), pp. 293–297. doi: 10.2307/4086771.

Piatt, J. F. and Nettleship, D. N. (1987) ‘Incidental catch of marine birds and mammals in fishing nets off Newfoundland, Canada’, *Marine Pollution Bulletin*, 18(6), pp. 344–349. doi: 10.1016/S0025-326X(87)80023-1.

Pimm, S., Raven, P., Peterson, A., Sekercioglu, C. H. and Ehrlich, P. R. (2006) ‘Human impacts on the rates of recent, present, and future bird extinctions.’, *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 103(29), pp. 10941–6. doi: 10.1073/pnas.0604181103.

Poinar, H., Kuch, M., McDonald, G., Martin, P. and Pääbo, S. (2003) ‘Nuclear Gene Sequences from a Late Pleistocene Sloth Coprolite’, *Current Biology*, 13(13), pp. 1150–1152. doi: 10.1016/S0960-9822(03)00450-0.

Poinar, H. N., Schwarz, C., Qi, J., Shapiro, B., Macphee, R. D. E., Buigues, B., Tikhonov, A., Huson, D. H., Tomsho, L. P., Auch, A., Rampp, M., Miller, W. and Schuster, S. C. (2006) ‘Metagenomics to Paleogenomics: Large-Scale Sequencing of Mammoth DNA’, *Science*, 311(5759), pp. 392–394. doi: 10.1126/science.1123360.

Pope, P. E. (2009) ‘Early Migratory Fishermen and Newfoundland’s Seabird Colonies’, *Journal of the North Atlantic*. Eagle Hill Foundation , 102, pp. 57–70. doi: 10.3721/037.002.s107.

Prost, S. and Anderson, C. N. K. (2011) ‘TempNet: a method to display statistical parsimony networks for heterochronous DNA sequence data’, *Methods in Ecology and Evolution*. Blackwell Publishing Ltd, 2(6), pp. 663–667. doi: 10.1111/j.2041-210X.2011.00129.x.

Prost, S., Smirnov, N., Fedorov, V. B., Sommer, R. S., Stiller, M., Nagel, D., Knapp, M. and Hofreiter, M. (2010) 'Influence of Climate Warming on Arctic Mammals? New Insights from Ancient DNA Studies of the Collared Lemming *Dicrostonyx torquatus*', *PLoS ONE*. Israel Program for Scientific Translations, 5(5), p. e10447. doi: 10.1371/journal.pone.0010447.

Purvis, A., Jones, K. E. and Mace, G. M. (2000) 'Extinction', *BioEssays*, 22(12), pp. 1123–1133. doi: 10.1002/1521-1878(200012)22:12<1123::AID-BIES10>3.0.CO;2-C.

Rambaut, A. (2000) 'Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies.', *Bioinformatics (Oxford, England)*, 16(4), pp. 395–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10869038 (Accessed: 2 February 2018).

Rambaut, A. (2016) *FigTree*. Available at: http://tree.bio.ed.ac.uk/software/figtree/.

Rambaut, A., Lam, T. T., Max Carvalho, L. and Pybus, O. G. (2016) 'Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen).', *Virus evolution*. Oxford University Press, 2(1), p. vew007. doi: 10.1093/ve/vew007.

Rambaut, A., Suchard, M. and Drummond, A. (2009) 'Tracer- MCMC Trace Analysis Tools'. Available at: http://tree.bio.ed.ac.uk/software/tracer/.

Ramsden, C., Melo, F. L., Figueiredo, L. M., Holmes, E. C., Zanotto, P. M. A. and VGDN Consortium (2008) 'High Rates of Molecular Evolution in Hantaviruses', *Molecular Biology and Evolution*, 25(7), pp. 1488–1492. doi: 10.1093/molbev/msn093.

Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J. S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., Gupta, R., Bertalan, M., Nielsen, K., Gilbert, M. T. P., Wang, Y., Raghavan, M., Campos, P. F., Kamp, H. M., Wilson, A. S., Gledhill, A., Tridico, S., Bunce, M., Lorenzen, E. D., Binladen, J., Guo, X., Zhao, J., Zhang, X., Zhang, H., Li, Z., Chen, M., Orlando, L., Kristiansen, K., Bak, M., Tommerup, N., Bendixen, C., Pierre, T. L., Grønnow, B., Meldgaard, M., Andreasen, C., Fedorova, S. A., Osipova, L. P., Higham, T. F. G., Ramsey, C. B., Hansen, T. V. O., Nielsen, F. C., Crawford, M. H., Brunak, S., Sicheritz-Pontén, T., Villems, R., Nielsen, R., Krogh, A., Wang, J. and Willerslev, E. (2010) 'Ancient human genome sequence of an extinct Palaeo-Eskimo.', *Nature*. NIH Public Access, 463(7282), pp. 757–762. doi: 10.1038/nature08835.

Raup, D. M. (1981) 'Extinction: bad genes or bad luck?', *Acta Geologica Hispanica*, 16(1–2), pp. 25–33.

Rawlence, N. J., Metcalf, J. L., Wood, J. R., Worthy, T. H., Austin, J. J. and Cooper, A. (2012) 'The effect of climate and environmental change on the megafaunal moa of New Zealand in the absence of humans', *Quaternary Science Reviews*, 50, pp. 141–153. doi: 10.1016/j.quascirev.2012.07.004.

Rawlence, N. J., Wood, J. R., Armstrong, K. N. and Cooper, A. (2009) 'DNA content and distribution in ancient feathers and potential to reconstruct the plumage of extinct avian taxa', *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1672). doi: 10.1098/rspb.2009.0755.

Rawlence, N., Kennedy, M., Waters, J. and Scofield, R. (2014) 'Morphological and ancient DNA analyses reveal inaccurate labels on two of Buller's bird specimens', *Journal of the Royal Society of New Zealand*. Taylor & Francis, 44(4), pp. 163–169. doi: 10.1080/03036758.2014.972962.

Reich, D., Green, R. E., Kircher, M., Krause, J., Patterson, N., Durand, E. Y., Viola, B., Briggs, A. W., Stenzel, U., Johnson, P. L. F., Maricic, T., Good, J. M., Marques-Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E. E., Stoneking, M., Richards, M., Talamo, S., Shunkov, M. V., Derevianko, A. P., Hublin, J.-J., Kelso, J., Slatkin, M. and Pääbo, S. (2010) 'Genetic history of an archaic hominin group from Denisova Cave in Siberia', *Nature*. Nature Research, 468(7327), pp. 1053–1060. doi: 10.1038/nature09710.

Richmond, D. J., Sinding, M.-H. S. and Gilbert, M. T. P. (2016) 'The potential and pitfalls of de-extinction', *Zoologica Scripta*, 45(S1), pp. 22–36. doi: 10.1111/zsc.12212.

Rizzi, E., Lari, M., Gigli, E., De Bellis, G., Caramelli, D., Magnussen, K., Steinmann, K., Kapranov, P., Thompson, J., Zazula, G., Froese, D., Moltke, I., Shapiro, B., Hofreiter, M., Al-Rasheid, K., Gilbert, M., Willerslev, E., Murphy, W., Götherström, A., Schuster, S., Bunce, M., Lorenzen, E., Binladen, J., Guo, X., Zhao, J., Zhang, X., Zhang, H., Li, Z., Chen, M., Orlando, L., Kristiansen, K., Bak, M., Tommerup, N., Bendixen, C., Pierre, T., Grønnow, B., Meldgaard, M., Andreasen, C., Fedorova, S., Osipova, L., Higham, T., Ramsey, C., Hansen, T., Nielsen, F., Crawford, M., Brunak, S., Sicheritz-Pontén, T., Villems, R., Nielsen, R., Krogh, A., Wang, J., Willerslev, E., Kelso, J., Lachmann, M., Reich, D. and Pääbo, S. (2012) 'Ancient DNA studies: new perspectives on old samples', *Genetics Selection Evolution*. BioMed Central, 44(1), p. 21. doi: 10.1186/1297-9686-44-21.

Rohland, N. and Hofreiter, M. (2007) 'Ancient DNA extraction from bones and teeth', *Nature Protocols*, 2(7), pp. 1756–1762. doi: 10.1038/nprot.2007.247.

Rohland, N., Siedel, H. and Hofreiter, M. (2004) 'Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens.', *BioTechniques*, 36(5), pp. 814–6, 818–21.

Rohland, N., Siedel, H. and Hofreiter, M. (2009) 'A rapid column-based ancient DNA extraction method for increased sample throughput', *Molecular Ecology Resources*. Blackwell Publishing Ltd, 10(4), pp. 677–683. doi: 10.1111/j.1755-0998.2009.02824.x.

De Santo, T. L. and Nelson, S. K. (1995) 'Chapter 3: Comparative Reproductive Ecology of the Auks (Family Alcidae) with Emphasis on the Marbled Murrelet', in Ralph, C. John; Hunt, George L., Jr.; Raphael, Martin G.; Piatt, J. F. (ed.) *Ecology and conservation of the Marbled Murrelet*. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, pp. 33–47.

Der Sarkissian, C., Allentoft, M. E., Avila-Arcos, M. C., Barnett, R., Campos, P. F., Cappellini, E., Ermini, L., Fernandez, R., da Fonseca, R., Ginolhac, A., Hansen, A. J., Jonsson, H., Korneliussen, T., Margaryan, A., Martin, M. D., Moreno-Mayar, J. V., Raghavan, M., Rasmussen, M., Velasco, M. S., Schroeder, H., Schubert, M., Seguin-Orlando, A., Wales, N., Gilbert, M. T. P., Willerslev, E. and Orlando, L. (2015) 'Ancient genomics', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1660), p. 20130387. doi: 10.1098/rstb.2013.0387.

Schippers, P., Stienen, E. W. M., Schotman, A. G. M., Snep, R. P. H. and Slim, P. A. (2011) 'The consequences of being colonial: Allee effects in metapopulations of seabirds', *Ecological Modelling*. Elsevier, 222(17), pp. 3061–3070. doi: 10.1016/J.ECOLMODEL.2011.05.022.

Schroeder, H., Ávila-Arcos, M. C., Malaspinas, A.-S., Poznik, G. D., Sandoval-Velasco, M., Carpenter, M. L., Moreno-Mayar, J. V., Sikora, M., Johnson, P. L. F., Allentoft, M. E., Samaniego, J. A., Haviser, J. B., Dee, M. W., Stafford, T. W., Salas, A., Orlando, L., Willerslev, E., Bustamante, C. D. and Gilbert, M. T. P. (2015) 'Genome-wide ancestry of 17th-century enslaved Africans from the Caribbean.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 112(12), pp. 3669–3673. doi: 10.1073/pnas.1421784112.

Schubert, M., Ermini, L., Sarkissian, C., Jonsson, H., Ginolhac, A., Schaefer, R., Martin, M. D., Fernandez, R., Kircher, M. and McCue, M. (2014) 'Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX', *Nat Protoc*, 9. doi: 10.1038/nprot.2014.063.

Schubert, M., Lindgreen, S. and Orlando, L. (2016) 'AdapterRemoval v2: rapid adapter trimming, identification, and read merging', *BMC Research Notes*, 9(1), p. 88. doi: 10.1186/s13104-016-1900-2.

Sekercioglu, Ç. H., Primack, R. B. and Wormworth, J. (2012) 'The effects of climate change on tropical birds', *Biological Conservation*, 148(1), pp. 1–18. doi: 10.1016/j.biocon.2011.10.019.

Sekercioglu, C. H., Schneider, S. H., Fay, J. P. and Loarie, S. R. (2008) 'Climate Change, Elevational Range Shifts, and Bird Extinctions', *Conservation Biology*, 22(1), pp. 140–150. doi: 10.1111/j.1523-1739.2007.00852.x.

Serjeantson, D. (2001) 'The great auk and the gannet: a prehistoric perspective on the extinction of the great auk', *International Journal of Osteoarchaeology*. John Wiley & Sons, Ltd., 11(1–2), pp. 43–55. doi: 10.1002/oa.545.

Shapiro, B., Drummond, A. J., Rambaut, A., Wilson, M. C., Matheus, P. E., Sher, A. V, Pybus, O. G., Gilbert, M. T. P., Barnes, I., Binladen, J., Willerslev, E., Hansen, A. J., Baryshnikov, G. F., Burns, J. A., Davydov, S., Driver, J. C., Froese, D. G., Harington, C. R., Keddie, G., Kosintsev, P., Kunz, M. L., Martin, L. D., Stephenson, R. O., Storer, J., Tedford, R., Zimov, S. and Cooper, A. (2004) 'Rise and Fall of the Beringian Steppe Bison', *Science*, 306(5701), pp. 1561–1565. doi: 10.1126/science.1101074.

Shapiro, B., Ho, S. Y. W., Drummond, A. J., Suchard, M. A., Pybus, O. G. and Rambaut, A. (2011) 'A Bayesian Phylogenetic Method to Estimate Unknown Sequence Ages', *Molecular Biology and Evolution*, 28(2), pp. 879–887. doi: 10.1093/molbev/msq262.

Shapiro, B., Sibthorpe, D., Rambaut, A., Austin, J., Wragg, G. M., Bininda-Emonds, O. R. P., Lee, P. L. M. and Cooper, A. (2002) 'Flight of the dodo.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 295(5560), p. 1683. doi: 10.1126/science.295.5560.1683.

Shepherd, L. D. (2017) 'A non-destructive DNA sampling technique for herbarium specimens', *PLOS ONE*. Edited by El. Public Library of Science, 12(8), p. e0183555. doi: 10.1371/journal.pone.0183555.

Shepherd, L., Tennyson, A. and Lambert, D. (2013) 'Using ancient DNA to enhance museum collections: a case study of rare kiwi (*Apteryx* spp.) specimens', *Journal of the Royal Society of New Zealand*. Taylor & Francis Group , 43(3), pp. 119–127. doi: 10.1080/03036758.2012.732585.

Shields, G. F. and Helm-Bychowski, K. M. (1988) 'Mitochondrial DNA of Birds', in R.F, J. (ed.) *Current Ornithology, vol 5*. Boston, MA: Springer US, pp. 273–295. doi: 10.1007/978-1-4615-6787-5_6.

Smith, A. N. (2015) 'Evolution of body mass in the Pan-Alcidae (Aves, Charadriiformes): the effects of combining neontological and paleontological data', *Paleobiology*, 42(1), pp. 8–26. doi: 10.5061/dryad.3k7v7.

Smith, C. I., Chamberlain, A. T., Riley, M. S., Cooper, A., Stringer, C. B. and Collins, M. J. (2001) 'Neanderthal DNA: Not just old but old and cold?', *Nature*. Nature Publishing Group, 410(6830), pp. 771–772. doi: 10.1038/35071177.

Smith, N. A. and Clarke, J. A. (2015) 'Systematics and evolution of the Pan-Alcidae (Aves, Charadriiformes)', *Journal of Avian Biology*. Blackwell Publishing Ltd, 46(2), pp. 125–140. doi: 10.1111/jav.00487.

Stamatakis, A. (2014) 'RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies', *Bioinformatics*, 30(9), pp. 1312–1313. doi: 10.1093/bioinformatics/btu033.

Stanton, J. C. (2014) 'Present-day risk assessment would have predicted the extinction of the passenger pigeon (*Ectopistes migratorius*)', *Biological Conservation*. Elsevier, 180, pp. 11–20. doi: 10.1016/J.BIOCON.2014.09.023.

Steenstrup, J. J. (1855) 'Et bidrag til Geirfuglens, *Alca impennis*, naturhistorie, og særligt til Kundskaben om dens tidligere udbredningskreds', *Vid. Meddr. Nat.hist. Foren*, pp. 33–118.

Stephens, P. A., Sutherland, W. J. and Freckleton, R. P. (1999) 'What Is the Allee Effect?', *Oikos*. Wiley Nordic Society Oikos, 87(1), p. 185. doi: 10.2307/3547011.

Stewart, J. R. (1999) ‘Intraspecific Variation in Modern and Quaternary European Lagopus’, *Smithsonian Contributions to Paleobiology*, 89, pp. 159–168.

Stiller, M., Baryshnikov, G., Bocherens, H., Grandal d’Anglade, A., Hilpert, B., Munzel, S. C., Pinhasi, R., Rabeder, G., Rosendahl, W., Trinkaus, E., Hofreiter, M. and Knapp, M. (2010) ‘Withering Away--25,000 Years of Genetic Decline Preceded Cave Bear Extinction’, *Molecular Biology and Evolution*, 27(5), pp. 975–978. doi: 10.1093/molbev/msq083.

Taberlet, P. and Bouvet, J. (1991) ‘A Single Plucked Feather as a Source of DNA for Bird Genetic Studies’, *The Auk*, 108(4), pp. 959–960.

Teasdale, M. D., Fiddyment, S., Vnouček, J., Mattiangeli, V., Speller, C., Binois, A., Carver, M., Dand, C., Newfield, T. P., Webb, C. C., Bradley, D. G. and Collins, M. J. (2017) ‘The York Gospels: a one thousand year biological palimpsest’, *bioRxiv*. doi: https://doi.org/10.1101/146324.

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F. N., de Siqueira, M. F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A. S., Midgley, G. F., Miles, L., Ortega-Huerta, M. A., Townsend Peterson, A., Phillips, O. L. and Williams, S. E. (2004) ‘Extinction risk from climate change’, *Nature*. Nature Publishing Group, 427(6970), pp. 145–148. doi: 10.1038/nature02121.

Thomas, J. E., Carvalho, G. R., Haile, J., Martin, M. D., Castruita, J. A. S., Niemann, J., Sinding, M.-H. S., Sandoval-Velasco, M., Rawlence, N. J., Fuller, E., Fjeldså, J., Hofreiter, M., Stewart, J. R., Gilbert, M. T. P. and Knapp, M. (2017) ‘An ʻAukward’ Tale: A Genetic Approach to Discover the Whereabouts of the Last Great Auks’, *Genes*. Multidisciplinary Digital Publishing Institute, 8(6), p. 164. doi: 10.3390/genes8060164.

Tigano, A., Damus, M., Birt, T. P., Morris-Pocock, J. A., Artukhin, Y. B., Friesen, V. L., VL, F., GJ, R., AJ, G. and RA, P. (2015) ‘The Arctic: Glacial Refugium or Area of Secondary Contact? Inference from the Population Genetic Structure of the Thick-Billed Murre (*Uria lomvia*), with Implications for Management’, *Journal of Heredity*. Arnold, London, 106(3), pp. 238–246. doi: 10.1093/jhered/esv016.

Tito, R. Y., Knights, D., Metcalf, J., Obregon-Tito, A. J., Cleeland, L., Najar, F., Roe, B., Reinhard, K., Sobolik, K., Belknap, S., Foster, M., Spicer, P., Knight, R. and Lewis, C. M. (2012) ‘Insights from Characterizing Extinct Human Gut Microbiomes’, *PLoS ONE*. Edited by M. Hofreiter. Public Library of Science, 7(12), p. e51146. doi: 10.1371/journal.pone.0051146.

Tito, R. Y., Macmil, S., Wiley, G., Najar, F., Cleeland, L., Qu, C., Wang, P., Romagne, F., Leonard, S., Ruiz, A. J., Reinhard, K., Roe, B. A. and Lewis, C. M. (2008) ‘Phylotyping and Functional Analysis of Two Ancient Human Microbiomes’, *PLoS ONE*. Public Library of Science, 3(11), p. e3703. doi: 10.1371/journal.pone.0003703.

To, T.-H., Jung, M., Lycett, S. and Gascuel, O. (2016) 'Fast Dating Using Least-Squares Criteria and Algorithms', *Systematic Biology*, 65(1), pp. 82–97. doi: 10.1093/sysbio/syv068.

Wales, N., Carøe, C., Sandoval-Velasco, M., Gamba, C., Barnett, R., Samaniego, J. A., Madrigal, J. R., Orlando, L. and Gilbert, M. T. P. (2015) 'New insights on single-stranded versus double-stranded DNA library preparation for ancient DNA', *BioTechniques*, 59(6), pp. 368–71. doi: 10.2144/000114364.

Warinner, C., Rodrigues, J. F. M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R. Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H. U., Salazar-García, D. C., Eppler, E., Seiler, R., Hansen, L. H., Castruita, J. A. S., Barkow-Oesterreicher, S., Teoh, K. Y., Kelstrup, C. D., Olsen, J. V, Nanni, P., Kawai, T., Willerslev, E., von Mering, C., Lewis, C. M., Collins, M. J., Gilbert, M. T. P., Rühli, F. and Cappellini, E. (2014) 'Pathogens and host immunity in the ancient human oral cavity', *Nature Genetics*. Nature Research, 46(4), pp. 336–344. doi: 10.1038/ng.2906.

Warinner, C., Speller, C. and Collins, M. J. (2014) 'A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1660), p. 20130376. doi: 10.1098/rstb.2013.0376.

Waters, J. M., Fraser, C. I., Maxwell, J. J. and Rawlence, N. J. (2017) 'Did interaction between human pressure and Little Ice Age drive biological turnover in New Zealand?', *Journal of Biogeography*, 44(7), pp. 1481–1490. doi: 10.1111/jbi.12991.

Wilhelm, S. I., Gilliland, S. G., Robertson, G. J., Ryan, P. C. and Elliot, R. D. (2008) 'Development and Validation of a Wing Key to Improve Harvest Management of Alcids in the Northwest Atlantic', *Journal of Wildlife Management*. Blackwell Publishing Ltd, 72(4), pp. 1026–1034. doi: 10.2193/2007-232.

Willerslev, E. and Cooper, A. (2005) 'Ancient DNA', *Proc. R. Soc. B*, 272(1558), pp. 3–16. doi: 10.1098/rspb.2004.2813.

Wojczulanis-Jakubas, K., Jakubas, D., Welcker, J., Harding, A. M. A., Karnovsky, N. J., Kidawa, D., Steen, H., Stempniewicz, L. and Camphuysen, C. J. (2011) 'Body size variation of a high-Arctic seabird: the dovekie (*Alle alle*)', *Polar Biology*. Springer-Verlag, 34(6), pp. 847–854. doi: 10.1007/s00300-010-0941-6.

Wojczulanis-Jakubas, K., Kilikowska, A., Fort, J., Gavrilo, M., Jakubas, D. and Friesen, V. (2015) 'No evidence of divergence at neutral genetic markers between the two morphologically different subspecies of the most numerous Arctic seabird', *Ibis*, 157(4), pp. 787–797. doi: 10.1111/ibi.12294.

Wojczulanis-Jakubas, K., Kilikowska, A., Harding, A. M. A., Jakubas, D., Karnovsky, N. J., Steen, H., Strøm, H., Welcker, J., Gavrilo, M., Lifjeld, J. T. and Johnsen, A. (2014) 'Weak population genetic differentiation in the most numerous Arctic seabird, the little auk', *Polar Biology*. Springer Berlin Heidelberg, 37(5), pp. 621–630. doi: 10.1007/s00300-014-1462-5.

Woodward, Weyand, N. and Bunnell, M. (1994) ‘DNA sequence from Cretaceous period bone fragments’, *Science*, 266(5188), p. 1229–1232. doi: 10.1126/science.7973705.

Wroe, S., Field, J. and Grayson, D. K. (2006) ‘Megafaunal extinction: climate, humans and assumptions.’, *Trends in ecology & evolution*. Elsevier, 21(2), pp. 61-2-4. doi: 10.1016/j.tree.2005.11.012.

Xiang, H., Gao, J., Yu, B., Zhou, H., Cai, D., Zhang, Y., Chen, X., Wang, X., Hofreiter, M. and Zhao, X. (2014) ‘Early Holocene chicken domestication in northern China.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 111(49), pp. 17564–17569. doi: 10.1073/pnas.1411882111.

Xie, W., Lewis, P. O., Fan, Y., Kuo, L. and Chen, M.-H. (2011) ‘Improving Marginal Likelihood Estimation for Bayesian Phylogenetic Model Selection’, *Systematic Biology*, 60(2), pp. 150–160. doi: 10.1093/sysbio/syq085.

Zhang, D.-X. and Hewitt, G. M. (2003) ‘Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects’, *Molecular Ecology*. Blackwell Science Ltd, 12(3), pp. 563–584. doi: 10.1046/j.1365-294X.2003.01773.x.

Zhang, G., Li, B., Li, C., Gilbert, M. T. P., Jarvis, E. D. and Wang, J. (2014) ‘Comparative genomic data of the Avian Phylogenomics Project’, *GigaScience*. Penn State University, Computer Science and Engineering, 3(1), p. 26. doi: 10.1186/2047-217X-3-26.

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., Storz, J. F., Antunes, A., Greenwold, M. J., Meredith, R. W., Ödeen, A., Cui, J., Zhou, Q., Xu, L., Pan, H., Wang, Z., Jin, L., Zhang, P., Hu, H., Yang, W., Hu, J., Xiao, J., Yang, Z., Liu, Y., Xie, Q., Yu, H., Lian, J., Wen, P., Zhang, F., Li, H., Zeng, Y., Xiong, Z., Liu, S., Zhou, L., Huang, Z., An, N., Wang, J., Zheng, Q., Xiong, Y., Wang, G., Wang, B., Wang, J., Fan, Y., da Fonseca, R. R., Alfaro-Núñez, A., Schubert, M., Orlando, L., Mourier, T., Howard, J. T., Ganapathy, G., Pfenning, A., Whitney, O., Rivas, M. V, Hara, E., Smith, J., Farré, M., Narayan, J., Slavov, G., Romanov, M. N., Borges, R., Machado, J. P., Khan, I., Springer, M. S., Gatesy, J., Hoffmann, F. G., Opazo, J. C., Håstad, O., Sawyer, R. H., Kim, H., Kim, K.-W., Kim, H. J., Cho, S., Li, N., Huang, Y., Bruford, M. W., Zhan, X., Dixon, A., Bertelsen, M. F., Derryberry, E., Warren, W., Wilson, R. K., Li, S., Ray, D. A., Green, R. E., O’Brien, S. J., Griffin, D., Johnson, W. E., Haussler, D., Ryder, O. A., Willerslev, E., Graves, G. R., Alström, P., Fjeldså, J., Mindell, D. P., Edwards, S. V, Braun, E. L., Rahbek, C., Burt, D. W., Houde, P., Zhang, Y., Yang, H., Wang, J., Avian Genome Consortium, A. G., Jarvis, E. D., Gilbert, M. T. P. and Wang, J. (2014) ‘Comparative genomics reveals insights into avian genome evolution and adaptation.’, *Science*. American Association for the Advancement of Science, 346(6215), pp. 1311–20. doi: 10.1126/science.1251385.