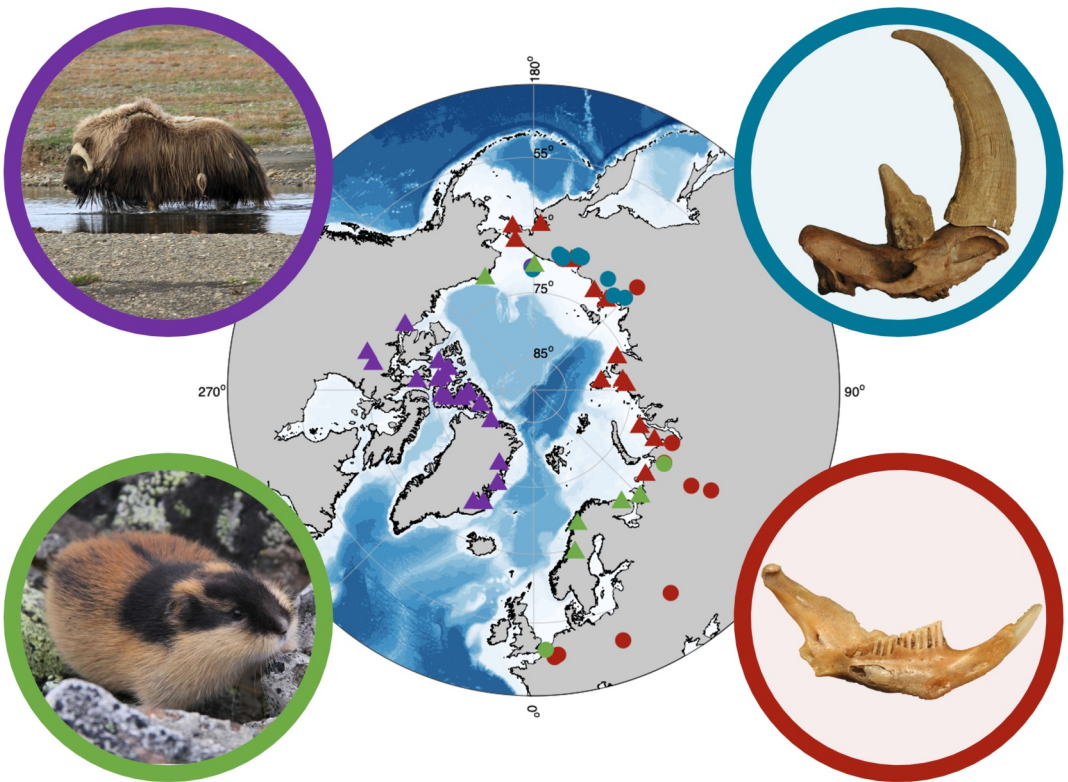


Investigating the impacts of Late Pleistocene climate change on Arctic mammals using palaeogenomics

Edana Lord



Investigating the impacts of Late Pleistocene climate change on Arctic mammals using palaeogenomics

Edana Lord

Academic dissertation for the Degree of Doctor of Philosophy in Systematic Zoology at Stockholm University to be publicly defended on Friday 21 October 2022 at 13.00 in Vivi Täckholmsalen (Q-salen), NPQ-huset, Svante Arrhenius väg 20.

Abstract

The climatic fluctuations of the Late Pleistocene likely had a large impact on the evolutionary history of Arctic species. Palaeogenomics is a useful tool to shed light on how past populations responded to these climatic shifts and the associated ice sheet dynamics and sea level change. Here, I have used modern and ancient DNA data from four Arctic mammals in order to investigate the impacts of Late Pleistocene climate on their evolutionary histories, from population dynamics and demography, to speciation and gene flow, adaptation, and genome erosion. In Paper I, using ancient mitogenomes from across their Late Pleistocene range, I showed that the Eurasian collared lemming (*Dicrostonyx torquatus*) had a dynamic Late Pleistocene population structure in Europe. Furthermore, the Eemian interglacial likely led to a bottleneck in collared lemmings, after which the species diversified during the Last Glacial period. Nuclear genome data from a modern individual in northeastern Siberia suggests population stability in northeastern Siberia during the Holocene. In Paper II, I sequenced the nuclear genome of a ~18,500 year old woolly rhinoceros (*Coelodonta antiquitatis*) and used this in combination with mitochondrial data to explore the demographic history of the species. There was little geographic structuring in the northeast Siberian population, and stability in their effective population size just prior to extinction, which may indicate a subsequent rapid decline towards extinction, likely associated with the Bølling-Allerød interstadial. Additionally, I found that this species had mutations in *TRPA1*, a gene involved in temperature sensitivity. In a third study (Paper III), I used whole genome data from modern and ancient true lemmings (*Lemmus* sp.) to determine that the Norwegian lemming (*L. lemmus*) has one of the youngest speciation times (~37-34 ka BP) of mammals. Norwegian lemmings have mutations in genes involved in coat colour, colour perception, fat transport and reproduction, and likely evolved their unique colouration as a result of isolation after the recolonisation of Fennoscandia. Finally, we examined the consequences of long-term small effective population size in muskox (*Ovibos moschatus*) using 107 modern nuclear genomes and one 21,000 year old Siberian genome (Paper IV). While muskox survived the warming at the end of the Late Pleistocene, the successive founder events experienced during its colonisation of the Canadian Arctic and Greenland reduced the genetic diversity to some of the lowest values observed in mammals. However, the results suggest that the long-term small population size likely led to purging of strongly deleterious alleles in the muskox, allowing them to persist to today with limited evidence of inbreeding depression. From a technical point, this thesis presents four de-novo genome assemblies, and the first whole nuclear genomes for these Arctic species. Taken together, the results in this thesis show that the climatic fluctuations, in particular the Eemian interglacial and Bølling-Allerød interstadial, along with sea level change and the formation and retreat of ice sheets during the Last Glacial Maximum have influenced the evolutionary histories of these four Arctic mammals.

Keywords: *palaeogenomics, Arctic, Late Pleistocene, past climate, collared lemming, woolly rhinoceros, Norwegian lemming, muskox, demography, population structure, speciation, adaptation, genome erosion.*

Stockholm 2022

<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-208657>

ISBN 978-91-7911-994-2
ISBN 978-91-7911-995-9



Stockholm
University

Department of Zoology

Stockholm University, 106 91 Stockholm

INVESTIGATING THE IMPACTS OF LATE PLEISTOCENE CLIMATE CHANGE ON ARCTIC MAMMALS USING PALAEOGENOMICS

Edana Lord

Investigating the impacts of Late Pleistocene climate change on Arctic mammals using palaeogenomics

Edana Lord

©Edana Lord, Stockholm University 2022

ISBN print 978-91-7911-994-2

ISBN PDF 978-91-7911-995-9

Front cover image by Edana Lord, photos used with permission.

Muskox, Wrangel Island - Love Dalén

Lemmus, Sarek, Sweden - Love Dalén

Woolly rhinoceros skull - from the personal archive of Fedor Shidlovskiy

Collared lemming mandible, Yangana Pe, Siberia - Eleftheria Palkopoulou

Map constructed in R v3.6.1 using the 'PlotSvalbard' package by Mikko Vihtakari

Printed in Sweden by Universitetsservice US-AB, Stockholm 2022

The thesis is based on the following articles, which are referred to in the text by their Roman numerals:

I **Lord, E.***, Marangoni, A.*, Baca, M., Popović, D., Goropashnaya, A. V., Stewart, J. R., Knul, M. V., Noiret, P., Germonpré, M., Jimenez, E-L., Abramson, N. I., Vartanyan, S., Prost, S., Smirnov, N. G., Kuzmina, E. A., Olsen, R-A., Fedorov, V. B., Dalén, L. Population dynamics and demographic history of Eurasian collared lemmings. Revised manuscript submitted.

II **Lord, E.***, Dussex, N.*, Kierczak, M., Díez-del-Molino, D., Ryder, O. A., Stanton, D. W. G., Gilbert, M. T. P., Sánchez-Barreiro, F., Zhang, G., Sinding, M.-H. S., Lorenzen, E. D., Willerslev, E., Protopov, A., Shidlovsky, F., Federov, S., Bocherens, H., Nathen, S. K. S. S., Goossens, B., van der Plicht, J., Chan, Y. L., Prost, S., Potapova, O., Kirillova, I., Lister, A. M., Heintzman, P. D., Kapp, J. D., Shapiro, B., Vartanyan, S., Götherström, A., Dalén, L. (2020). Pre-extinction demographic stability and genomic signatures of adaptation in the woolly rhinoceros. *Current Biology*, 30, 1–9.

III **Lord, E.**, Feinauer, I. S., Lagerholm, V. K., Näsval, K., Ersmark, E., Olsen, R-A., Prost, S., Kuzmina, E. A., Smirnov, N. G., Stewart, J. R., Noiret, P., Germonpré, M., Ehrich, D., Ims, R., Berglund, H., Fedorov, V. B., Goropashnaya, A. V., Dalén, L., Díez-del-Molino, D. Genome analyses suggest recent speciation, rapid adaptation, and post-glacial isolation in the Norwegian Lemming. Manuscript.

IV Pečnerová, P., **Lord, E.**, Garcia-Erill, G., Hanghøj, K., Rasmussen, M. S., Meisner, J., Liu, X., Santander, C. G., Hvilsom, C., Götherström, A., Carøe, C., Dalerum, F., Måsviken, J., Vartanyan, S., Heide-Jørgensen, M. P., Sinding, M.-H. S., Schmidt, N. M., Dalén, L., Albrechtsen, A., Heller, R., Moltke, I., Siegismund, H. Population genomics of the muskox’ resilience in the near absence of genetic variation. Manuscript.

*Shared first authors

Candidate contributions to thesis articles*

	I	II	III	IV
Conceived the study	Substantial	Significant	Minor	Significant
Designed the study	Substantial	Significant	Significant	Significant
Collected the data	Substantial	Substantial	Minor	Significant
Analysed the data	Substantial	Substantial	Substantial	Significant
Manuscript preparation	Substantial	Substantial	Substantial	Significant

***Contribution Explanation**

Minor: contributed in some way, but contribution was limited.
Significant: provided a significant contribution to the work.
Substantial: took the lead role and performed the majority of the work.

I have also co-authored the following articles, which are not included in the thesis:

Kutschera, V. E., Kierczak, M., van der Valk, T., von Seth, J., Dussex, N., **Lord, E.**, Dehasque, M., Stanton, D.W.G., Emami, P., Nystedt, B., Dalén, L., Díez-del-Molino, D. (2022) GenErode: a bioinformatics pipeline to investigate genome erosion in endangered and extinct species. *BMC Bioinformatics*. 23(1), 1-17.

Dussex, N., Bergfeldt, N., de Anca Prado, V., Dehasque, M., Díez-del-Molino, D., Ersmark, E., Kanellidou, F., Larsson, P., Lemež, S., **Lord, E.**, Mármol-Sánchez, E., Meleg, I. N., Måsviken, J., Naidoo, T., Studerus, J., Vicente, M., von Seth, J., Götherström, A., Dalén, L., Heintzman, P. D. (2021) Integrating multi-taxon palaeogenomes and sedimentary ancient DNA to study past ecosystem dynamics. *Proceedings of the Royal Society B*, 288(1957), 20211252.

Rosengren, E., Acatrinei, A., Cruceru, N., Dehasque, M., Haliuc, A., **Lord, E.**, Mircea, C.I., Rusu, I., Mármol-Sánchez, E., Kelemen, B.S., Meleg, I.N. (2021) Ancient faunal history revealed by interdisciplinary biomolecular approaches. *Diversity*, 13(8), 370.

Gosling, A. L., **Lord, E.**, Boocock, J., Cameron-Christie, S., Horsburgh, K. A., Kardailsky, O., Prost, S., Wilcox, S., Addison, D., Thompson, A., Kalolo, J., the Genographic Consortium, Matisoo-Smith, E. A. (2021) A Population history of Tokelau – genetic variation and change in atoll populations. *Journal of Island and Coastal Archaeology*, 11-18.

Lord, E., Collins, C., deFrance, S., LeFebvre, M. J., Pigièrre, F., Eeckhout, P., Erauw, C., Fitzpatrick, S. M., Healy, P. F., Martínez-Polanco, M. F., Garcia, J. L., Ramos Roca, E., Delgado, M., Sánchez Urriago, A., Peña León, G. A., Toyne, J. M., Dahlstedt, A., Moore, K. M., Laguer Diaz, C., Zori, C., Matisoo-Smith, E. (2020) Ancient DNA of Guinea Pigs (*Cavia* spp.) Indicates a Probable New Center of Domestication and Pathways of Global Distribution. *Scientific Reports*, 10(1), 8901.

Table of contents

Introduction	11
Late Pleistocene climate	11
Palaeogenomics	13
Cold-adapted Arctic species	13
Aims	15
Methods	15
Sample collection	15
DNA extraction, library preparation, and sequencing	15
Reference genome assembly	17
Bioinformatic processing of resequencing data	17
Population structure, divergence, and gene flow	18
Reconstructing demographic history	18
Genetic diversity, inbreeding and genetic load	19
Adaptation	19
Results	19
Datasets	19
Phylogenetic structure and divergence	20
Demographic histories of Arctic species	23
Genomic basis of adaptation	23
Genome erosion	26
Discussion	28
Impact of climate on Late Pleistocene population dynamics and demography	28
Speciation and divergence	30
Adaptation in the woolly rhinoceros and Norwegian lemmings	31
Genetic consequences of small population size	31
Impact of human predation on megafauna	32
Generating reference genome assemblies for non-model taxa	32
Future directions for palaeogenomics and past climate change	33
Implications for Arctic species under anthropogenic climate warming	34
Conclusions	34
References	36
Acknowledgements	41
Svensk sammanfattning	43

Introduction

Late Pleistocene climate

The Late Pleistocene was characterised by climatic fluctuations, which led to changes in species' distributions, in turn influencing the evolutionary history of species (Hewitt, 1996, 2000). The climate varied between long cold glacial phases such as the Last Glacial period, which occurred 115-11.7 thousand years before present (ka BP) and included the Last Glacial Maximum (LGM, 26.5-19 ka BP), and short, warm interglacials including the Eemian (130-115 ka BP) and the Holocene (11.7 ka BP to present) (Fig 1). Furthermore, there were also shorter warm interstadials (e.g. the Bølling-Allerød, 14.7-12.8 ka BP) and cold stadial phases (e.g. the Younger-Dryas, 12.8-11.7 ka BP) during the glacial phases (Svensson *et al.*, 2006). During cooler periods, the tundra environment extended southwards, encompassing large parts of Eurasia and North America. During these phases, the fossil record suggests that cold-adapted Arctic species such as woolly mammoths (*Mammuthus primigenius*), woolly rhinoceroses (*Coelodonta antiquitatis*), muskox (*Ovibos moschatus*) and collared lemmings (*Dicrostonyx torquatus*) were distributed across Eurasia (Markova *et al.*, 1995; Lister and Sher, 2001; Lorenzen *et al.*, 2011; Stuart and Lister, 2012). Expansions during cooler phases likely led to diversification of Arctic species, whereas during interglacial and possibly also interstadial phases, Arctic species inhabited smaller geographical distributions, termed polar refugia (Stewart *et al.*, 2010). The range contractions during warm periods likely had a major impact on the demographic history of Arctic species, including reductions in population size, the loss of genetic diversity, as well as population and species extinctions (Lister and Stuart, 2008). Consequently, the warming after the LGM and during the Holocene saw the extinction of a number of cold-adapted taxa, including the woolly rhinoceros, cave lions (*Panthera spelaea*), cave bears (*Ursus spelaeus*), and woolly mammoths (Stuart and Lister, 2007, 2012; Dehasque *et al.*, 2021).

A consequence of the climatic fluctuations was the formation of large ice sheets during glacial phases. During the Last Glacial period, ice sheets covered Scandinavia (Fennoscandian ice sheet) and also a large part of North America (Laurentide ice sheet). These ice sheets reached their maximum extent between 26.5-19 ka BP (Clark *et al.*, 2009) and 26-25 ka BP (Stokes, 2017), respectively. Ice sheets restricted the movement of species to glacier-free areas during the LGM. Conversely, during the warming at the end of the Pleistocene, the retreat of the ice sheets enabled the recolonisation of previously glaciated areas. Both the formation and subsequent retreat of the ice sheets allowed for opportunities for speciation and divergence within species. In Scandinavia, the Norwegian lemming (*Lemmus lemmus*), brown bear (*Ursus arctos*), and moose (*Alces alces*), along with various plant species including spruce (*Picea abies*), recolonised the area after the LGM (Lagerholm *et al.*, 2014; Ersmark *et al.*, 2019; Dussex *et al.*, 2020; Nota *et al.*, 2022). However, the source populations for the recolonisation of Fennoscandia are largely unknown for several taxa, with origins suggested in northern Europe for moose (Dussex *et al.*, 2020; Meiri *et al.*, 2020), and Siberia or a local refugium in Scandinavia for Norwegian lemmings (Lagerholm *et al.*, 2014). In North America, Hansen *et*

al., (2018) suggested that the two subspecies of muskox arose as a result of separation into northern and southern populations by the Laurentide ice sheet.

The formation and melting of the ice sheets contributed to dramatic changes in sea levels between glacial and interglacial phases, respectively. During the LGM, the sea level was up to 120 m lower than today given the water present in the large ice sheets. This exposed a larger land area in the Northern Hemisphere, and allowed for movement over land bridges, for example between Eurasia and North America via the Bering land bridge. This likely facilitated the dispersal of species and gene flow between lineages that were isolated during interglacial phases. Wolves (*Canis lupus*), horses (*Equus* sp.), muskox, moose and woolly mammoths all dispersed via this land bridge at varying times in the Late Pleistocene (Campos *et al.*, 2010; Palkopoulou *et al.*, 2013; Meiri *et al.*, 2020; Vershinina *et al.*, 2021; Bergström *et al.*, 2022).

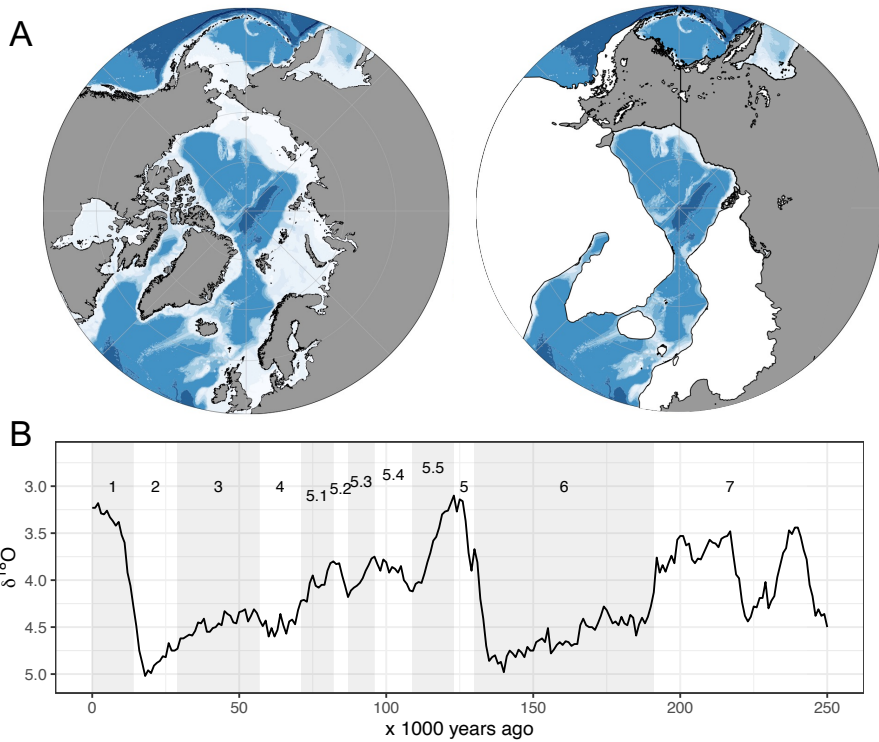


Figure 1. Late Pleistocene climate. A) Land area, and extent of ice sheets in the Northern Hemisphere during interglacials (left) and the Last Glacial Maximum (right). The Laurentide Ice sheet is depicted over North America, and the Fennoscandian ice sheet over Scandinavia and Europe. B) Changes in oxygen isotope ratio in the last 250,000 years (proxy for temperature record), with Marine Isotope Stages (MIS) listed. MIS 5-5.5 indicates the Eemian interglacial. MIS 2 corresponds to the Last Glacial Maximum. MIS 1 is the Holocene interglacial. Maps were constructed in R v4.0.0 using ‘ggOceanMaps’ (Vihtakari, 2022) and data from Batchelor *et al.*, (2019) and Assis *et al.*, (2018). The climate curve was constructed in Rv4.0.0 using the ‘gsloid’ package and data from Lisecki and Raymo (2005) and Ruan *et al.*, (2016).

Additionally, as the sea level was elevated during interglacials, populations became isolated, for example woolly mammoths on Wrangel and St Paul's Islands (Graham *et al.*, 2016). For woolly mammoths, isolation on islands also led to sustained small population sizes, which in turn led to inbreeding and loss of genetic diversity (Palkopoulou *et al.*, 2015; Pečnerová *et al.*, 2017), and ultimately extinction.

Palaeogenomics

Palaeogenomics is a powerful tool for studying the past. It can reveal information about the evolutionary history of a species, and how they interacted with other species and their environment. Ancient DNA can be extracted from various materials, including: bone, teeth, hair, soft tissues, calculus, faeces, sediment, and other organic material. Early ancient DNA studies focused on using Polymerase Chain Reaction (PCR) to amplify specific regions of the genome, for example the mitochondrial cytochrome B gene or hypervariable region. However, with the development of second generation sequencing technologies, and increased computing and bioinformatics infrastructure in the past decade, recovering and analysing whole ancient nuclear genomes is now much more feasible (Orlando *et al.*, 2013; Palkopoulou *et al.*, 2015), including from samples over one million years old (van der Valk *et al.*, 2021). This increase in the number of loci allows for investigating the evolutionary history, population dynamics, gene flow, adaptation, and changes in allele frequencies over time (see Mitchell and Rawlence, 2021). Furthermore, high-throughput sequencing and targeted capture methods can be used to identify and study ancient diseases (Spyrou *et al.*, 2019), and reconstruct ancient ecosystems from sediments (Willerslev *et al.*, 2003; Edwards, 2020).

Palaeogenetic studies have been undertaken on a range of extinct and extant Arctic taxa, owing to better DNA preservation in northern latitudes. Demographic histories of Arctic taxa have begun to be explored, including woolly mammoths, muskox, bison (*Bison* sp.), and the woolly rhinoceros (Campos *et al.*, 2010; Lorenzen *et al.*, 2011; Palkopoulou *et al.*, 2015). Ancient DNA has indicated that there was higher genetic diversity and population structure in Late Pleistocene lemmings, muskox, wolves, and arctic foxes (*Vulpes lagopus*) (Campos *et al.*, 2010; Brace *et al.*, 2012; Palkopoulou *et al.*, 2016; Larsson *et al.*, 2019). Following this, the role of genome erosion (defined in this thesis as the loss of genetic diversity, and increased inbreeding and mutational load) in the extinction process has been explored in the woolly mammoth (Palkopoulou *et al.*, 2015; Pečnerová *et al.*, 2017; Rogers and Slatkin, 2017). Furthermore, recent studies have examined gene flow, for example in bears (Barlow *et al.*, 2018; Cahill *et al.*, 2018), that occurred when the species' ranges overlapped in the past. Most studies to-date have focused on single species histories, with the exception of sedimentary ancient DNA studies that tend to focus on multiple taxa (e.g., Gelabert *et al.*, 2021).

Cold-adapted Arctic species

The high latitudes of the Arctic were home to a number of cold-adapted species during the Late Pleistocene. While some went extinct, including the woolly rhinoceros, others have survived the warming at the end of the Pleistocene and inhabit the Arctic tundra today, although in a

vastly reduced distribution. This thesis focused on four Arctic species: Eurasian collared lemmings, woolly rhinoceros, Norwegian lemmings, and muskox. These Arctic mammals were focused on because of the large changes in their geographic distribution exhibited during the Late Pleistocene climatic fluctuations. During glacial periods, these species were found in Eurasia, from the United Kingdom to Siberia, and also in North America for muskox. However, during interglacial periods, they were restricted to refugia in the northernmost parts of their range. While studies have looked at the impact of past climate on megafauna, it can be difficult to discern between the effects of climate and human predation (Lorenzen *et al.*, 2011). However, small mammals that were likely less affected by human predation, such as lemmings, provide an opportunity to study the effect of climate alone. In addition, for palaeogenetic studies, material from higher and colder latitudes tends to preserve DNA better over longer periods of time, and each of these species is well-characterised in the fossil record. Ancient DNA studies on cold-adapted Arctic species to-date have primarily focused on short mitochondrial regions, with ancient whole nuclear genome data currently only available for mammoths (Palkopoulou *et al.*, 2015; van der Valk *et al.*, 2021), grey wolves (Bergström *et al.*, 2022), polar bears (*Ursus maritimus*) and brown bears (Cahill *et al.*, 2018), and moose (Dussex *et al.*, 2020).

Eurasian collared lemmings are small cold-adapted rodents that inhabit the tundra of Siberia. The fossil record indicates this species was widespread across Eurasia during the Late Pleistocene (Markova *et al.*, 1995). Previous mitochondrial DNA studies suggest that this species underwent population extinction and replacement events coinciding with climatic shifts (Prost *et al.*, 2010; Brace *et al.*, 2012; Palkopoulou *et al.*, 2016). Today, the populations are separated into distinct geographical clusters (Fedorov *et al.*, 2020).

The woolly rhinoceros was a megaherbivore that inhabited Eurasia during the Pleistocene. This species had a similar distribution to the woolly mammoth, although they did not cross the Bering land bridge to North America (Stuart and Lister, 2012). Previous ancient DNA research suggested that this species underwent a demographic expansion during the LGM (Lorenzen *et al.*, 2011). The woolly rhinoceros went extinct ~14 ka BP after a range contraction towards northeastern Siberia, with both climate and overhunting proposed as possible extinction factors (Lorenzen *et al.*, 2011; Stuart and Lister, 2012).

True lemmings (*Lemmus* sp.) inhabit the holarctic tundra today. The Norwegian lemming is the only mammal endemic to Fennoscandia. This species recolonised Fennoscandia after the LGM, although the origins of this recolonisation are still unclear (Fedorov and Stenseth, 2001; Lagerholm *et al.*, 2014). The Norwegian lemming is known for its bright colouration and aggressive anti-predatory behaviour. Previous research has shown it is a sister species to the Siberian brown lemming (*L. sibiricus*) (Fedorov *et al.*, 1999), although phylogenetic relationships within the Siberian true lemmings remain unclear.

Unlike woolly rhinoceros, the muskox is a megaherbivore that survived the warming during the transition to the Holocene. There are two subspecies of muskox today, the barren-ground muskox in the Canadian mainland, and the white-faced muskox in the Canadian Arctic and

Greenland. Muskox colonised these areas during a northward expansion of the species in the Holocene, reaching Greenland at ~5.4 ka (Bennike and Andreassen, 2005). Previous genomic data suggested that the muskox has very low genetic diversity (Hansen *et al.*, 2018), likely as a result of founder events during colonisation of the Arctic.

Aims

The focus of this thesis was to use ancient and modern mitochondrial and nuclear genomes to understand the impacts of Late Pleistocene climate change on four Arctic taxa. More specifically, this thesis aimed to investigate (a) Late Pleistocene population dynamics and demographic history, (b) divergence within and between species, (c) species-specific genomic adaptations, and (d) genome erosion in response to environmental pressures. *Paper I* explored changes in the population history of collared lemmings during the Late Pleistocene. *Paper II* focused on the demographic history and genomic adaptations of the extinct woolly rhinoceros. *Paper III* evaluated the recent speciation and role of post-glacial isolation in the Norwegian lemming. Finally, *Paper IV* investigated the demographic history and consequences of long-term small population size on genome erosion in muskox.

Methods

Sample collection

Samples for *Papers I-IV* were collected from palaeontological and/or archaeological sites, or from opportunistic surface collections during expeditions. Sampling locations ranged across the Holarctic and Europe (Fig 2). For *Paper I*, 59 Eurasian collared lemming samples (>50-0.5 ka BP) were collected from across their Late Pleistocene range, and one *Dicrostonyx* sp. sample from the Batagaika permafrost crater in Siberia. In *Paper II*, 14 woolly rhinoceroses were sampled from northeastern Siberia, dating to 50-14 ka BP. For *Paper III*, true lemming samples were collected from four modern taxa: the Norwegian lemming, the Siberian brown lemming (eastern and western lineages) and the Nearctic brown lemming. Additionally, two ancient *Lemmus* sp. from the United Kingdom (12.5 ka BP) and Russia (9.6 ka BP) were included. Finally for *Paper IV*, 107 modern muskox samples were collected from their native range across Canada and Greenland, along with a single 21 ka old sample from Wrangel Island. A subset of samples were radiocarbon dated and calibrated to calendar years with Oxcal 4.4 (Ramsey, 2009).

DNA extraction, library preparation, and sequencing

The majority of the ancient DNA lab work was performed at either the Swedish Museum of Natural History, or the Centre of Palaeogenetics, both in Stockholm. Surfaces of the bone were cleaned using a Dremel drill prior to bone powder being collected by drilling into the bone. For *Paper I*, given the small size of the samples, exogenous contamination was removed by crushing a subset of the bone and then performing a bleach (~0.5%) wash and pre-digestion

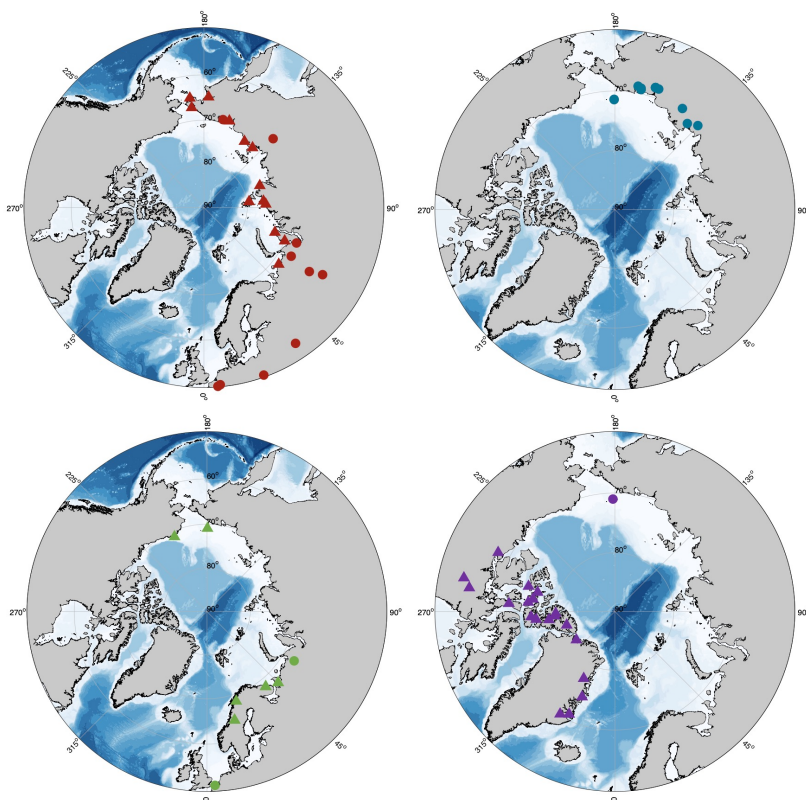


Figure 2. Maps of sample sites for each of the four Arctic mammals included in this thesis. Eurasian collared lemmings (red), woolly rhinoceros (blue), true lemmings (green), and muskox (purple) are shown. Circles indicate ancient samples, and triangles represent modern samples. The maps were produced using the ‘plotSvalbard’ package in Rv3.6.1 (Vihtakari, 2020).

with the extraction buffer. For all papers, ancient DNA was extracted using Protocol C (Yang *et al.*, 1998), following modifications in Dehasque *et al.*, (2022). Double-stranded DNA libraries compatible with Illumina sequencing platforms were prepared from the DNA extracts following Meyer and Kircher (2010). Libraries were subject to an initial round of sequencing to identify samples with sufficient endogenous content to proceed to further sequencing (mitogenome or whole genome). For samples that subsequently underwent whole genome resequencing, additional DNA libraries were prepared to maximise the complexity of the sample. DNA sequencing of the Illumina-prepared libraries was performed at the Science for Life Laboratory (SciLifeLab) in Stockholm using Illumina sequencing technology. For a subset of samples in *Paper II*, DNA libraries were prepared using the BEST2 method (Carøe *et al.*, 2018). These libraries were sequenced at the Beijing Genomics Institute (BGI), China. All modern samples were extracted from blood or muscle tissue using the Kingfisher Cell and Tissue DNA kit on a Kingfisher DNA extraction robot (Thermo Fisher Scientific). Modern DNA libraries were constructed using the TruSeq PCR-free protocol (Illumina) and sequenced at SciLifeLab, Sweden.

Reference genome assembly

To construct a reference genome, a high-quality DNA extract from modern fresh/frozen tissue is required. For *Paper I* and *III*, tissue was collected, and modern DNA extracted as described above. DNA libraries of differing lengths (short insert and mate-pair) were sequenced at SciLifeLab in Stockholm, Sweden. Three genome assembly methods were tested: SOAPdenovo (Li *et al.*, 2010), ALLPATHS (Butler *et al.*, 2008) and ABySS (Simpson *et al.*, 2009) and evaluated using BUSCO (Seppey, Manni and Zdobnov, 2019). In both cases, ALLPATHS produced the most contiguous and complete assembly. In *Paper II*, the assembly for Sumatran rhinoceros (*Dicerorhinus sumatrensis*), was generated using Chicago, Hi-C, mate-pair and short read libraries. The assembly was conducted using the HiRise pipeline (Dovetail Genomics). For *Paper IV*, the muskox genome assembly was conducted at BGI using the 10x Genomics pipeline.

Bioinformatic processing of resequencing data

All samples in *Papers I-III* and the ancient sample in *Paper IV* were processed from raw data to BAM files (mitogenome samples) or VCFs (high coverage nuclear samples) in a development version of the GenErode pipeline (<https://github.com/NBISweden/GenErode>; Kutschera *et al.*, 2022). This pipeline is designed to process ancient/historical and modern samples using the specific settings required for each data type, from raw reads to VCF, and conduct basic population structure analyses (principal component analysis, PCA) and several genome erosion metrics (genome wide heterozygosity, runs of homozygosity, and relative mutational load). While GenErode is primarily designed to be used with high coverage data (> 6X genome-wide coverage), the processing steps from raw reads to BAM files can be used for low coverage data as well. Briefly, GenErode implements a pipeline to trim and align raw reads to a reference genome, here using the reference genomes generated in this thesis or outgroup reference genomes acquired from NCBI. The development version of GenErode predominantly used in this thesis differs from the published version in that a modified version of SeqPrep (<https://github.com/jstjohn/SeqPrep>) was used for trimming and merging of ancient FASTQ data, as per Palkopoulou *et al.*, (2015), and Trimmomatic v2 was used for trimming the adapters from modern FASTQ data (Bolger, Lohse and Usadel, 2014). The modern samples in *Paper IV* were analysed using a second snakemake pipeline from Pečnerová *et al.*, (2021).

For the mitochondrial analyses, trimmed reads were aligned to the respective mitochondrial reference genome for each species using BWA aln with specific settings for ancient DNA: deactivated seeding (-l 16,500), allowing more substitutions (-n 0.01), and allowing up to two gaps (-o 2). Duplicates were removed using a custom Python script. Coverage was estimated using SAMtools and plotted with a custom R script. Consensus FASTA files were generated using Geneious with a majority rule and at least 3X coverage required to call positions (Kearse *et al.*, 2012). All ambiguous sites were called as N. One mitogenome in *Paper I* was generated using mitoBIM (Hahn, Bachmann and Chevreux, 2013).

Population structure, divergence, and gene flow

To explore how populations have changed over time and how different species related to each other, phylogenetic trees were constructed using both mitochondrial and nuclear data. Time-calibrated mitochondrial phylogenetic trees were constructed using BEAST v1.10.4 (Drummond and Rambaut, 2007). Nuclear genome trees were constructed in three ways. Firstly, for the datasets with high and low coverage nuclear genome data, trees were constructed using an identity-by-state (IBS) matrix, included in the ANGSD toolkit (Korneliussen, Albrechtsen and Nielsen, 2014), and a balanced minimum evolution or neighbour joining tree was constructed. To test for topology switching across the genome in *Paper III*, we evaluated the different topologies of 1 Mb trees using Twisst (Martin and Van Belleghem, 2017). Thirdly, in *Paper IV* we generated a time-calibrated nuclear phylogenetic tree with iqtree2 (To *et al.*, 2016). Population structure was also investigated using PCAnsd (Meisner and Albrechtsen, 2018) and ngsAdmix (Skotte, Korneliussen and Albrechtsen, 2013), and Estimation of Effective Migration Surfaces (EEMS) (Petkova, Novembre and Stephens, 2016) was used to examine barriers to gene flow.

Divergence times were taken from the time-calibrated phylogenetic trees for both the mitochondrial and nuclear datasets. Additionally *Paper III* estimated divergence times based on a method described in Palkopoulou *et al.*, (2015), where the demographic history of a pseudo-diploid X chromosome was constructed using the Pairwise Sequential Markovian Coalescent (PSMC) approach (Li and Durbin, 2011). During divergence, the coalescence rate is zero, thus the effective population size (taken as $1/\text{coalescence rate}$) trends towards infinity.

To test for the presence of gene flow in the datasets in *Papers III* and *IV*, we used the D-statistic, or ABBA-BABA (Green *et al.*, 2010; Durand *et al.*, 2011) test implemented in ANGSD or popstats (Skoglund *et al.*, 2015). We additionally used Treemix (Pickrell and Pritchard, 2012) and admixture graph modelling with qpGraph (Patterson *et al.*, 2012) in *Paper IV* to assess whether the phylogenetic structure of the dataset was better explained with gene flow between one or more populations.

Reconstructing demographic history

Using the mitochondrial datasets in *Papers I* and *II*, the demographic history of the Eurasian collared lemmings and the woolly rhinoceros, respectively, was examined using a Bayesian Skyline approach in BEAST v1.10.4 (Drummond and Rambaut, 2007). This approach estimates the female effective population size (N_{ef}) through time. Additionally in *Paper I*, the demographic history was explored using an Approximate Bayesian Computation approach with simulated data. The mitochondrial data was simulated using fastsimcoal v2.27.9 (Excoffier *et al.*, 2021), and four demographic models were tested.

The demographic history was reconstructed for each species in *Papers I*, *II* and *IV* using PSMC (Li and Durbin, 2011). This method uses the local density of heterozygous sites across the genome to assess changes in effective population size in the time to the most recent common

ancestor, where the coalescence rate is inversely proportional to the effective population size. Mutations rates and generation times taken from the literature were used to scale the curves.

Genetic diversity, inbreeding and genetic load

Characterising changes in genome-wide diversity can elucidate the effects of genetic bottlenecks, founder effects and changes in population size. Genome-wide heterozygosity was estimated using PLINK (Purcell *et al.*, 2007) for hard-called genotypes (*Papers I, II, IV*) and ANGSD for genotype-likelihoods (*Paper IV* only). Runs of homozygosity, that is stretches of the genome which are identical on both chromosomes, were estimated using PLINK and ROHan (Renaud *et al.*, 2019).

In *Paper IV*, relative mutational load was estimated using GERP++ scores (Davydov *et al.*, 2010), implemented in the GenErode pipeline. GERP scores are the number of mutations in conserved regions of the genome, here estimated using genomes across mammalian taxa. Having a higher rate of mutations in such areas is considered to have a deleterious effect on the species (Huber, Kim and Lohmueller, 2020). We additionally used SnpEff (Cingolani *et al.*, 2012) to determine the number of mutations in the high, moderate, and modifier, and low impact categories to investigate the impacts on realised and potential load. Here, realised load was calculated as the proportion of deleterious variants that are homozygous in individuals, whereas potential load was all the deleterious alleles present in both heterozygous and homozygous state (Mathur and DeWoody, 2021).

Adaptation

In *Papers II* and *III* we evaluated the genomic basis of adaptation. For this, we identified the non-synonymous mutations with SnpEff (Cingolani *et al.*, 2012). We then selected a focal set of genes with either mutations that led to the most significant change in amino acid physicochemical properties (*Paper II*) or the highest number of non-synonymous mutations, along with any genes with loss of function mutations (*Paper III*). Using the gene sets we carried out gene ontology enrichment analyses using Panther (Mi *et al.*, 2013) with the horse as a reference set (*Paper II*), or using GOrilla (Eden *et al.*, 2009) with the Norway rat as a reference set (*Paper III*). Finally, in *Paper II*, we compared our set of genes to that previously identified in the woolly mammoth (Lynch *et al.*, 2015).

Results

Datasets

For all papers, we generated *de-novo* reference genome assemblies that were used for mapping the short-read resequencing data. For *Paper I*, the genome of a Eurasian collared lemming from northeastern Siberia was assembled. In *Paper II*, we generated an assembly for the Sumatran rhinoceros, the closest extant species to the woolly rhinoceros. We generated reference genome assemblies for the Norwegian lemming and muskox for *Paper III* and *Paper IV*, respectively.

In *Paper I*, we reconstructed 59 ancient mitochondrial genomes (> 50-0.5 ka BP) from across the Late Pleistocene geographical distribution (Belgium to northeastern Siberia) of Eurasian collared lemming and analysed these along with 54 published modern mitogenomes. We additionally aligned the short-read data generated for the reference genome assembly against itself to analyse one high-coverage nuclear genome. The dataset for *Paper II* consisted of the first full nuclear genome of the woolly rhinoceros, generated from an 18,500 year old individual from Chukotka, Siberia. We also analysed 14 mitogenomes from northeastern Siberia, ranging in age from 50-14 ka BP. For *Paper III*, we resequenced 11 modern genomes from four true lemming taxa, including the Norwegian lemming, the Siberian brown lemming Western and Eastern lineages, and the Nearctic brown lemming (*L. trimucronatus*). In addition, we generated low-coverage nuclear data for two Late Pleistocene *Lemmus* sp. samples, one 12.5 ka BP from Bridged Pot Cave, United Kingdom, and the other 9.6 ka BP from Pymva Shor, Russia, as well as short mitochondrial sequences (522 bp) for 55 samples. Finally, in *Paper IV*, we analysed 107 modern nuclear genomes from across the muskox' native range, along with that of a 21 ka old Siberian muskox from Wrangel Island.

Phylogenetic structure and divergence

Paper I showed that the Eurasian collared lemmings had a distinct population structure during the Late Pleistocene, with five mitochondrial clades (Fig 3) similar to that previously observed from short mtDNA data (Brace *et al.*, 2012; Palkopoulou *et al.*, 2016). Only one clade is present today, with a most recent common female ancestor occurring ~25 ka BP, during the LGM. This clade shows geographical structuring, representing the different subpopulations in northern Siberia, along with a now-extinct subclade in Europe. Both the Eurasian and North American collared lemmings appear to have diversified after ~100 ka BP and ~111 ka BP, respectively. In contrast to the collared lemmings, *Paper II* found an absence of geographic structuring among the northeastern Siberian woolly rhinoceros samples investigated (Fig 4). Both major clades in our dataset overlapped geographically and were present up until the extinction of the species at ~14 ka BP.

We constructed both mitogenome and nuclear phylogenies for true lemmings and muskox. In true lemmings, these are concordant and both phylogenies show the separation of each species into monophyletic clades (Fig 5). The ancient *Lemmus* samples form a monophyletic clade with modern Norwegian lemmings. The mitogenome and nuclear pseudo-diploid X chromosome divergence times are consistent, showing that the Siberian brown lemming East and West lineages diverged ~133-124 ka BP and then the Siberian West lineage and Norwegian lemmings diverged ~37-34 ka BP. In addition to the complete concordance and high support in both phylogenies, we also did not find any indication of gene flow between the modern *Lemmus* lineages.

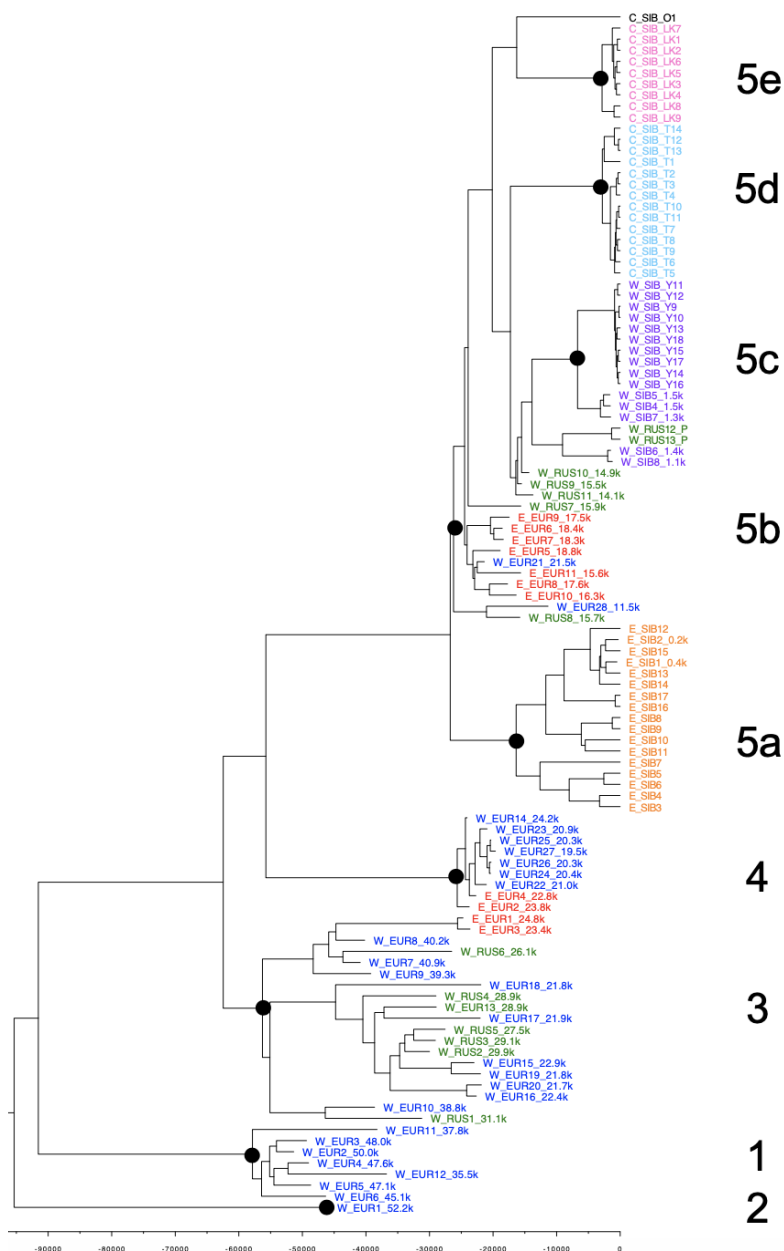


Figure 3. Mitochondrial phylogeny of Eurasian collared lemmings. Each major clade is numbered, and the most recent common ancestor is indicated with a black circle. The colours correspond to geographical regions: blue = western Europe (W_EUR), red = eastern Europe (E_EUR), green = western Russia (W_RUS), purple = western Siberia (W_SIB), light blue = central Siberia Taymyr region (C_SIB_T), light pink = central Siberia Yana-Kolyma region (C_SIB_LK), black = central Siberia Ostrov Bolshevik (C_SIB_O), orange = eastern Siberia (E_SIB).

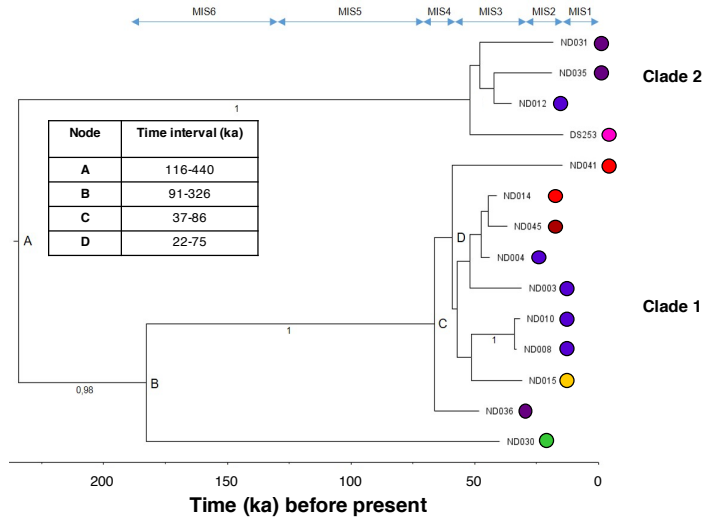


Figure 4. Mitogenome phylogeny of the woolly rhinoceros in northeastern Siberia (50-14 ka BP). Coloured dots represent the different localities sampled: yellow = Yana river, pink = Tumat, dark red = Abyysky, red = Kolyma river, blue = west Chukotka, purple = north Chukotka, and green = Wrangel Island. Divergence time ranges are shown for nodes A-D. Marine Isotope Stages (MIS) 1-6 are represented by arrows.

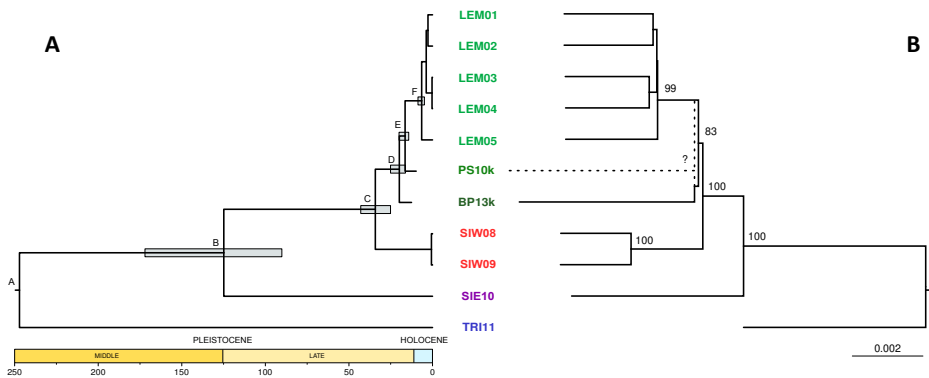


Figure 5. Mitogenome and nuclear phylogeny for true lemmings (*Lemmus* sp.). A) Time-calibrated mitogenome phylogeny depicting: green = Norwegian lemmings, dark green = ancient *Lemmus*, red = Siberian lemming western lineage, purple = Siberian lemming eastern lineage, blue = Nearctic brown lemming. Major nodes are labelled (A-F) and the 95% higher posterior density intervals for the divergence times are represented by grey bars. B) Nuclear genome phylogeny of true lemmings using a balanced minimum evolution tree based on an identity-by-state matrix. The approximate position of PS10k is indicated with a dashed line due to a lack of overlapping data between the two ancient *Lemmus* to build the tree with both samples. Numbers indicate bootstrap support for each node, and the scale bar demonstrates phylogenetic distance.

In muskox, the mitogenome and nuclear phylogenies both place the ancient Siberian sample basal to modern samples (Fig 6), although the structuring within present day muskox is inconsistent between the two phylogenies. Each geographic cluster forms a monophyletic clade in the nuclear phylogeny; however, the mitogenome phylogeny indicates that a subset of Canadian mainland samples form a divergent clade from the remaining modern samples. These remaining modern samples form two clades, which demonstrate geographical structuring. Additional population structure and admixture analyses in the muskox supported the geographic clustering seen in the nuclear phylogeny. However, these analyses also suggested some gene flow between lineages, with treemix results suggesting migration from the Canadian mainland to the Canadian Islands and from the Canadian Islands to the Greenland North population, and admixture graph modelling supporting only the first migration edge.

Demographic histories of Arctic species

The mitochondrial demographic modelling for collared lemmings supports a model with both an Eemian and a LGM bottleneck. *Paper 1* also showed a population decline during the LGM in the Bayesian Skyline analysis for this species, although that may reflect the extinction of the mitogenome clades in our dataset during the LGM. The demographic reconstruction for the woolly rhinoceros based on mitogenome data supports a constant population size through time.

We then reconstructed the Pleistocene demographic trajectories for Eurasian collared lemmings, woolly rhinoceroses, and muskox using PSMC (Fig 7). The demographic history of collared lemmings shows a peak in effective population size (N_e) occurring during the Last Glacial period at ~55-50 ka BP, after which N_e decreased until 10.2 ka BP. The effective population size then increased during the Holocene and remained relatively stable until the present day. In the woolly rhinoceros, N_e increased during the penultimate glacial period, and then gradually decreased from ~127 ka BP until the LGM. There was then a sharp increase in N_e , followed by stability until the death of the individual at 18.5 ka BP. Muskox had a relatively stable population size until ~30 ka BP, whereby the population size decreased gradually until ~5 ka BP. The mainland and island populations then maintained a stable population size at ~500 and ~300 individuals, respectively.

Genomic basis of adaptation

In both the woolly rhinoceros and Norwegian lemmings, we identified genes with non-synonymous mutations (i.e. those that change the amino acid and thus can affect the function of the protein). We found 1524 and 536 genes with non-synonymous mutations in woolly rhinoceros (n=1) and Norwegian lemmings (n=5), respectively. In the woolly rhinoceros, these genes are involved in multiple biological pathways including cellular component organisation or biogenesis, cellular processes, localization, reproduction, biological regulation, response to stimulus, developmental processes, and metabolic processes. We found 89 genes previously identified in the woolly mammoth (Lynch *et al.*, 2015) that also have non-synonymous mutations in the woolly rhinoceros. Interestingly, one of these genes is TRPA1 which is involved in temperature sensitivity.

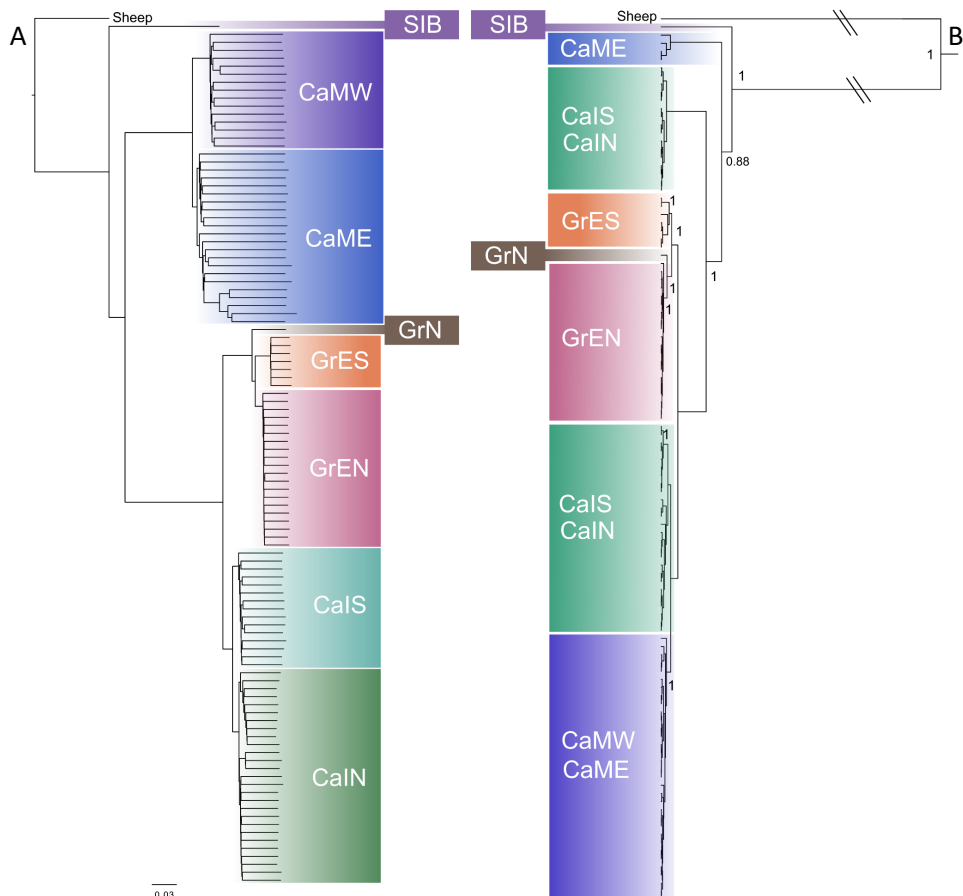


Figure 6. Nuclear and mitochondrial genome phylogenies of muskox. A) Nuclear genome neighbour-joining phylogeny constructed using an identity-by-state matrix, with the sheep as an outgroup. Populations are: light purple = Siberia (SIB), dark purple = West Canadian mainland (CaMW), blue = East Canadian mainland (CaME) brown = North Greenland (GrN), orange = East Greenland south (GrES), pink = East Greenland north (GrEN), light green = South Canadian islands (CaIS), dark green = North Canadian islands (CaIN). B) Mitogenome phylogeny constructed in BEAST with posterior probabilities on the branches. Population labelling follows A.

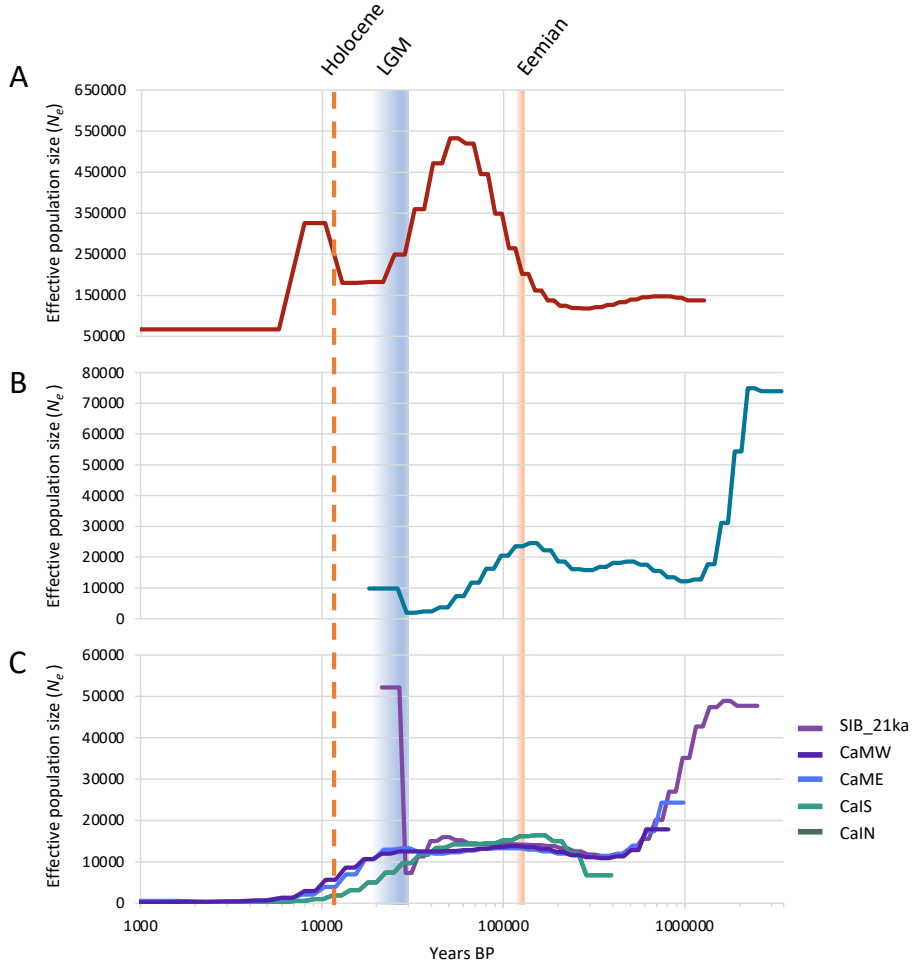


Figure 7. Demographic trajectories of Eurasian collared lemmings, the woolly rhinoceros, and muskox as inferred via PSMC. A) Collared lemmings PSMC scaled using a generation time of 2 generations/year and a mutation rate of 5×10^{-9} substitutions/site/generation. B) Woolly rhinoceros PSMC scaled using a generation time of 12 years and a mutation rate of 2.34×10^{-8} substitutions/site/generation. C) PSMC of ancient and modern muskox, using one individual per locality. Samples are coloured by population: light purple = Siberia (Sib, EL012), dark green = North Canadian islands (CaIN, AH1), light green = South Canadian islands (CaIS, AI1), blue = East Canadian mainland (CaME, BL10), and dark purple = West Canadian mainland (CaMW, KU11). The PSMC was scaled using a generation time of 8 years and a mutation rate of 1.00×10^{-8} substitutions/site/generation. Note that the effective population size limits on the y-axis differ for each species.

In the Norwegian lemming, we focused on 21 genes that are fixed for the derived allele in all five modern Norwegian lemming genomes and that had three or more non-synonymous mutations, including loss of function mutations. These genes were involved in pigmentation (*Myo1a*, *Lyst*), colour perception (*Atf6*), reproduction (*Fsip2*, *POTE-a*, *Plekha5*, *Spag17*), and fat transport (*Apob*, *Bco*, *Relch*). *Myo1a* may also play a role in thermoregulation (Jia *et al.*, 2019). We found that for the majority of the uniquely derived genes in the modern Norwegian lemmings, the two ancient samples shared the ancestral allele with the other True lemming species and an outgroup, the prairie vole (*Microtus ochrogaster*).

Genome erosion

In *Paper II*, we estimated the genome-wide heterozygosity of the woolly rhinoceros to 1.7 heterozygous sites per 1000 bp. This estimate is higher than that found in Late Pleistocene mainland and Wrangel Island mammoths (Palkopoulou *et al.*, 2015), and living rhinoceroses (Mays *et al.*, 2018; Tunstall *et al.*, 2018; Liu *et al.*, 2021; von Seth *et al.*, 2021). Furthermore, there was no indication of recent bottlenecks based on both the length and quantity of runs of homozygosity.

In the muskox, *Paper IV* explored changes in genetic diversity, inbreeding, and relative mutational load on both temporal and geographic scales (Fig 8). We found that the ancient Siberian muskox had higher diversity than any of the present-day populations. Furthermore, diversity decreased from south to north, with the Greenlandic samples demonstrating the lowest diversity. The white-faced muskox in the Canadian islands and Greenland have the second lowest genome-wide diversity observed in a mammal so far. In parallel, inbreeding increased from south to north, with Eastern Greenland samples having up to 80% of their genomes in runs of homozygosity. Relative mutational load showed a more varied pattern, where the Siberian individual had the highest load, and there was a reduction in load from south to north, until the northernmost Canadian Arctic populations. Interestingly, the Jameson Land samples in eastern Greenland had the lowest genome-wide heterozygosity, and the highest relative mutational load.

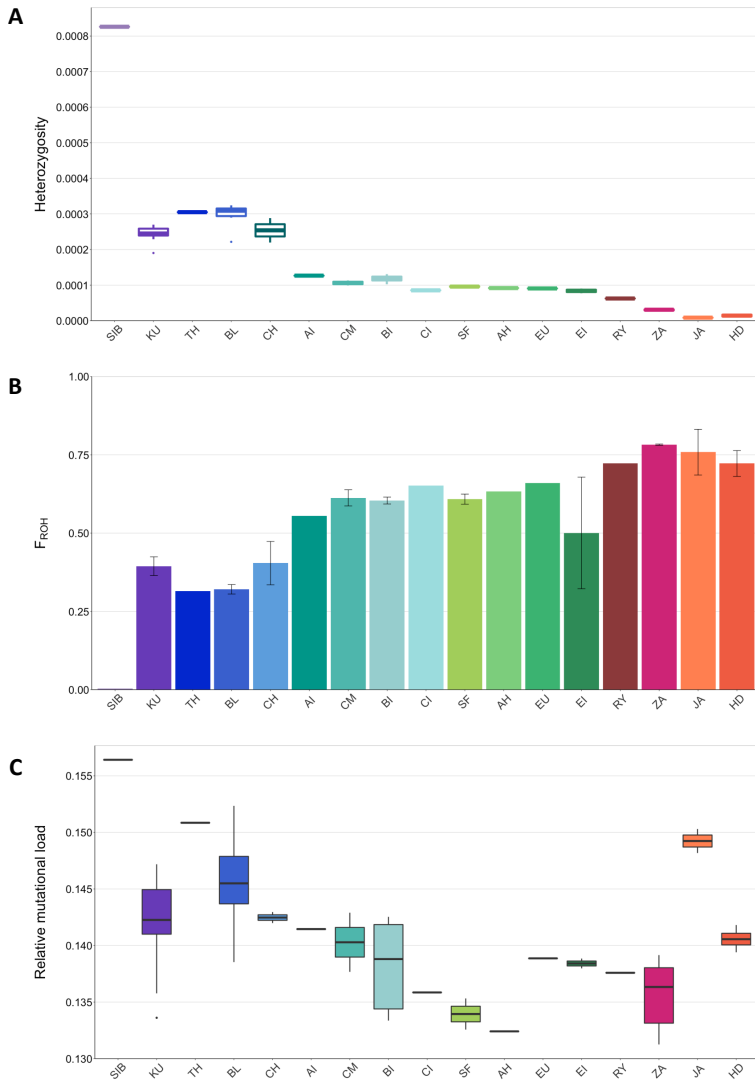


Figure 8. Genome-wide heterozygosity, inbreeding (F_{ROH}), and relative mutational load in muskox. Populations are listed in geographical order from eastern Siberia (SIB), West Canadian mainland (KU), East Canadian mainland (TH, BL, CH), South Canadian islands (AI, CM, BI, CI), North Canadian islands (SF, AH, EU, EI), North Greenland (RY), East Greenland north (ZA), East Greenland south (JA, HD).

Discussion

Impact of climate on Late Pleistocene population dynamics and demography

The Late Pleistocene climatic fluctuations had an impact on Arctic species' population structure and their demographic histories. As Arctic species inhabited large ranges across Eurasia during the Last Glacial period, it is not surprising that they had greater diversity in lineages in the past. *Paper I* showed that this is most striking in Eurasian collared lemmings, with five distinct mitochondrial lineages, and only one of which persists today (Fig 3). It is plausible that these Pleistocene lineages were the result of expansions during favourable climatic conditions, from source populations in Siberia or local refugia within Europe (Brace *et al.*, 2012; Palkopoulou *et al.*, 2016). Four of the five lineages appear to have gone extinct during the LGM, or by the onset of the Holocene. While this may be a result of incomplete sampling, it may also be due to replacement by a divergent lineage from eastern Siberia. Nuclear genomes of each of the four lineages will be required to test whether gene flow occurred amongst the Late Pleistocene lineages of collared lemmings.

Extinct European mitochondrial lineages were observed in Late Pleistocene true lemmings (*Paper III*, also see Lagerholm *et al.*, 2014) and have been documented previously for muskox (Campos *et al.*, 2010), woolly rhinoceroses (Lorenzen *et al.*, 2011; Rey-Iglesia *et al.*, 2021), cave lions (Barnett *et al.*, 2009), and mammoths (Nyström *et al.*, 2010; Palkopoulou *et al.*, 2013), among other species (Hofreiter *et al.*, 2004). The Late Pleistocene European lineages of true lemmings, collared lemmings, and muskox all became extinct during the transition to the Holocene, where rapid warming meant that Arctic species were unable to track their habitat northwards. Local extinctions may have also occurred during the short interstadials within the Last Glacial Period, although fine-scale temporal data from additional species would be needed to further explore this hypothesis.

Climate also played an important role in the deeper demographic histories of Arctic species, with population sizes varying in response to changes in range size and founder events. Within Eurasian collared lemmings, our mitochondrial demographic analyses support a bottleneck during the Eemian interglacial, consistent with the subsequent diversification of both the Eurasian and North American lineages during the early Last Glacial period. PSMC results suggested a decline during the last glacial period, likely due to the extinction of the four Late Pleistocene lineages and subsequent replacement of a divergent lineage from Siberia. This appears to be in contrast to the fossil record which indicates that collared lemmings were widespread throughout the last glacial period. However, it is important to note that N_e does not equate to census population size, and the decrease in N_e seen here may in fact represent the loss of genetic variation due to lineage extinction during the LGM.

Similarly to collared lemmings, N_e in the woolly rhinoceros declined during much of the Last Glacial period, with the onset of this decline occurring during the Eemian interglacial (Fig 7). However, we note that this peak could also be the result of population subdivision. Woolly rhinoceroses then experienced a rapid increase in N_e at the onset of the LGM. This increase is

consistent with previous mitochondrial data and the fossil record (Lorenzen *et al.*, 2011; Stuart and Lister, 2012). The population stability just prior to extinction for the woolly rhinoceros implies its population size likely declined rapidly towards extinction, which fossil data suggests occurred during the warming of the Bølling-Allerød interstadial (14.7-12.8 ka BP).

The PSMC analyses indicate population stability in northeastern Siberia for the woolly rhinoceros shortly prior to extinction and for Eurasian collared lemmings during the Holocene. Furthermore, both the woolly rhinoceros and the collared lemming genomes analysed here showed no signs of recent inbreeding, suggesting they had not undergone recent bottlenecks. Population stability in northeastern Siberia during this time period has also been observed in wolves (Loog *et al.*, 2020), as well as humans (Sikora *et al.*, 2019). This may indicate that a more stable environment was maintained in northeastern Siberia during the Late Pleistocene and transition to the Holocene. This is in contrast to the woolly mammoth, which experienced successive local extinctions as its range contracted northwards during this time (Dehasque *et al.*, 2021), suggesting general patterns may not pertain to all species.

Paper IV indicates that muskox had a relatively stable population size during much of the Late Pleistocene, however, the effective population size declined from ~30-5 ka BP. The beginning of this decline may have been associated with the formation of the Laurentide ice sheet, which reached the maximum extent at 26-25 ka BP (Stokes, 2017). The subsequent decrease in N_e was likely associated with successive founder events during the northward colonisation of the Canadian Arctic during the Holocene and ultimately Greenland by ~5.4 ka BP (Bennike and Andreassen, 2005).

The primary method used in this thesis to investigate long-term demographic changes was PSMC analysis. PSMC can be useful in that it only requires a single high-coverage (i.e. > 8X coverage) diploid genome to trace the demographic history of a species (Li and Durbin, 2011). However, it also applies a number of assumptions (e.g. population panmixia) that may be violated in natural populations, as demonstrated in the highly structured populations of collared lemmings, where population structure and population subdivision can lead to spurious peaks (i.e. inflated N_e) in the curve (Mazet *et al.*, 2016; Mather, Traves and Ho, 2020). It is also important to note that PSMC typically smoothes out abrupt changes in N_e , thus short time-scale bottlenecks are not usually observed (Li and Durbin, 2011). Furthermore, for non-model species, including extinct species, approximate estimations of mutation rate and generation time must be used. Changing one or both of these parameters will cause temporal shifts in the PSMC curve, making interpretations approximate, especially in regard to specific events such as environmental change or human presence. Other methods for modelling demography are available but usually require multiple high-coverage genomes (e.g. MSMC, $\delta a\delta i$, SMC++, SFS in fastsimcoal), making them less suitable for ancient genome projects with limited datasets.

Despite the caveats of the PSMC approach, this thesis demonstrates that the climatic fluctuations of the Late Pleistocene likely played a role in the demographic histories of multiple Arctic mammals. Demographic trajectories from Arctic species not investigated here may reveal broad patterns across taxa, but so far the data indicates that the Eemian interglacial likely

had an impact on both Eurasian collared lemmings and woolly rhinoceros. Furthermore, PSMC was able to capture the population decline associated with Holocene warming and colonisation of the Arctic for muskox.

Speciation and divergence

The climatic fluctuations appeared to have played a large role in the speciation of true lemmings. Firstly, the divergence between the two species of Siberian lemming, *L. sibiricus* West and East likely occurred during the Eemian interglacial. It could be hypothesised that *L. sibiricus* East was isolated on Wrangel Island during the Eemian interglacial and then recolonised northeastern Siberia during the Last Glacial period. However, further modern and ancient genomes from this species, especially across the northeastern part of mainland Siberia and from Wrangel Island, will be required to characterise the evolutionary history of the Siberian true lemmings.

Secondly, our results show that the Norwegian lemming and its sister taxa, *L. sibiricus* West diverged from a common ancestor just prior to the LGM at ~37-34 ka BP, demonstrating that this species has one of the youngest speciation times among mammals. The number of generations (34,000-37,000) is equal to or less than that between the polar and brown bear (34,000-48,000 generations), which is currently considered the most recent mammalian speciation event (Liu *et al.*, 2014; Hirata *et al.*, 2017). The Norwegian lemming may have become isolated from the Siberian lemming during a pre-LGM colonisation of Scandinavia (Larsen *et al.*, 1987) or at the onset of the LGM. Results from short mitochondrial data suggested that a Late Pleistocene *Lemmus* lineage that diverged from the Siberian lemming was distributed across Europe and western Russia during the last glaciation (Lagerholm *et al.*, 2014). As the Fennoscandian ice sheet expanded during the LGM, it has been hypothesised that a small population of this lineage became isolated in a northern ice-free refugium, and subsequently evolved into the Norwegian lemming (Fedorov and Stenseth, 2001; Lagerholm *et al.*, 2014). Our results support this hypothesis as the modern Norwegian lemmings formed a monophyletic clade and have uniquely derived variants that are not present in the ancient *Lemmus*. This indicates that most of the unique adaptations in the Norwegian lemming likely evolved within Fennoscandia. Additionally, given that we find no indication of Holocene gene flow with the Siberian lemming, the Norwegian lemming likely underwent post-glacial isolation. Thus, the LGM and subsequent retreat of the ice sheet was instrumental in the evolution of the Norwegian lemming.

The LGM also played a role in the divergence between the two subspecies of muskox. As the Laurentide ice sheet formed, previous genetic and fossil evidence suggests that the two populations became isolated, with one north and the other one south of the ice sheet (Harington, 1980; Lent, 1999; Lorenzen *et al.*, 2011; Hansen *et al.*, 2018). The mitochondrial divergence times estimated here between the two main modern lineages indicate a divergence time of ~19-10 ka BP, suggesting a complex history associated with the colonisation of the High Arctic. The extent to which each of the isolated populations contributed to present day populations

remains unknown. Recovery of ancient genomes from the localities in which muskox existed during the LGM would aid in further understanding the impacts of the LGM on this species.

Adaptation in the woolly rhinoceros and Norwegian lemmings

In this thesis, the genomic basis of adaptation was explored in two taxa: the woolly rhinoceros and the Norwegian lemming. In the woolly rhinoceros, we observed mutations in genes relating to temperature sensation, i.e. TRPA1. This is similar to what was previously observed in the mammoth (Lynch *et al.*, 2015). It is plausible that these mutations were involved in the ability of the woolly rhinoceros to withstand the extreme cold of Late Pleistocene Eurasia.

Norwegian lemmings display unique colouration and behaviour among the true lemmings. *Paper II* identified mutations in genes associated with pigmentation, perception of colour, fat transport and reproduction. It is possible these mutations led to the unique colouration and ability to recognise conspecifics in the Norwegian lemming. While our limited ancient data indicated that the two Late Pleistocene samples had the ancestral alleles for several of these mutations, future research with high-coverage ancient genomes will be required to determine whether the Norwegian lemming evolved its unique mutations prior to, or after the LGM.

Genetic consequences of small population size

Recent palaeogenomic evidence has suggested that mammoths lost genetic diversity and experienced increased inbreeding associated with their isolation on Wrangel Island and subsequent extinction (Palkopoulou *et al.*, 2015; Pečnerová *et al.*, 2017; Pečnerová, 2018). In contrast, the genome of the 18.5 ka BP woolly rhinoceros individual that we analysed had a higher genome-wide heterozygosity than mainland mammoths, no evidence of recent inbreeding, and revealed that woolly rhinoceroses had a stable effective population size just prior to extinction. In combination, this indicates that the woolly rhinoceros population in northeastern Siberia experienced no genomic erosion effects typical of small populations. This also implies a sudden and rapid population decline of woolly rhinoceros towards extinction, supported by the fossil record (Stuart and Lister, 2012). Whole genomes from samples closer to the extinction may show elevated levels of inbreeding and a decline in genome-wide heterozygosity.

While many megafaunal species went extinct during the later part of the Pleistocene and early Holocene, muskox persisted, albeit at a very low effective population size. The impact of isolation and post-glacial recolonisation on the diversity of Arctic fauna has been documented previously in the Arctic fox (Dalén *et al.*, 2005), Arctic hare (*Lepus arcticus*) (Waltari and Cook, 2005), grey wolves (Sinding *et al.*, 2018), and caribou (*Rangifer tarandus*) (Klüttsch *et al.*, 2017). Muskox experienced a series of founder events during the colonisation of the Canadian Arctic and Greenland in the Holocene (Harington, 1980; Lent, 1999; Hansen *et al.*, 2018). As a result, present-day muskox have some of the lowest genetic diversity observed in mammals with the white-faced muskox having the second lowest genetic diversity recorded to date in mammals. Low genetic diversity and effective population size often results in high

levels of inbreeding, which in turn exposes harmful genetic mutations due to increased rates of homozygosity (Hedrick, 2001). However, recent evidence from endangered populations has suggested that species may be able to purge these mutations under conditions of extremely low population size (van der Valk *et al.*, 2019; Dussex *et al.*, 2021; Morin *et al.*, 2021). Thus, it is likely that the long-term low effective population size of muskox contributed to purging highly deleterious mutations from the population, leading to a reduced genetic load in some of the modern populations (Fig 8). In the case of the Jameson Land population in East Greenland, we observe low diversity, high inbreeding, and high mutational load. Thus, it may be this population has experienced more recent inbreeding and/or purging has not occurred to the same extent as in the other Greenlandic populations. Ancient samples from before and after each successive colonisation event across the Canadian Arctic and Greenland will elucidate finer scale changes in diversity and the impact of the founder events in the muskox.

Impact of human predation on megafauna

Here, I have focused primarily on the impact of climate in the population histories of Arctic taxa. However, human predation may have also had an impact, particularly on the woolly rhinoceros and muskox, which are known to have been preyed by humans (Lorenzen *et al.*, 2011). In the woolly rhinoceros, we do not see a decrease in effective population size associated with the first appearance of anatomically modern humans in northeastern Siberia at 31.6 ka BP. However, there is currently limited knowledge of the changes of human population density in this region during the warming after the LGM. Increasing human population density during this time may have played a role in the extinction of woolly rhinoceros, however additional research into ancient human population dynamics from northeastern Siberia is required. Muskox on the other hand were known to be an important subsistence source for Arctic peoples, such as the Independence I culture (Knuth, 1967), and also were subject to intense hunting from the late 19th-early 20th century (Barr, 1991; Cuyler *et al.*, 2020). Demographic declines associated with recent human impacts, through hunting or introduction of predators, have been documented for a number of species, including Arctic fox (Larsson *et al.*, 2019). Thus, historical DNA studies will be useful to understand the demographic consequences of recent hunting practices in muskox.

Generating reference genome assemblies for non-model taxa

An aspect of this thesis that is important to highlight is that for each study species, we generated reference genome assemblies. Closely related reference genomes enable accurate characterisation of genetic diversity, population structure, and reconstruction of demographic histories of species. Recently large consortia have formed such as the Vertebrate Genomes Project (Rhie *et al.*, 2021), including the Bird 10,000 Genomes project (B10K; Feng *et al.*, 2020), and the Darwin Tree of Life project (<https://www.darwintreeoflife.org/>), with the aims of producing as many reference genomes and genome data as possible for non-model organisms. This will significantly aid genomic research on such species, enabling further understanding of the evolutionary history of non-model taxa. Genomic data may also play an important role in the future of conservation genetic practices, thus generating references will

enable studies across temporal and geographic scales for species currently under threat (Paez *et al.*, 2022).

Future directions for palaeogenomics and past climate change

This thesis highlighted an impact of the Eemian interglacial in the demographic history and diversification of Arctic species, in particular collared lemmings. Importantly, *Paper I* also demonstrated that for these small rodents, it was possible to retrieve high levels of endogenous ancient DNA from samples ~50 ka BP and up to 330 ka BP. Future ancient DNA studies could focus on samples before, during, and after the Eemian interglacial period, to fully characterise the impact of this interglacial on Arctic species. This would allow for an assessment of the extent to which Arctic species experienced bottlenecks during the Eemian and whether unique lineages were lost, similar to that observed during the transition to the Holocene (Campos *et al.*, 2010; Brace *et al.*, 2012).

From a geographical standpoint, northeastern Siberia appears to be a hotspot for genetic diversity, for both woolly rhinoceroses and Eurasian collared lemmings, along with horses, bison, and wolves (Kirillova *et al.*, 2015; Fages *et al.*, 2019; Fedorov *et al.*, 2020; Loog *et al.*, 2020). Further characterisation of this region during the Last Glacial period, in terms of plant and animal species composition, and their diversity, might indicate that the environment was more favourable for cold-adapted species. This could be achieved using sedimentary ancient DNA, which is becoming a prominent tool for reconstructing past landscapes (Voldstad *et al.*, 2020; Crump *et al.*, 2021). Furthermore, methods could be developed to infer within-species haplotypic diversity from sedimentary ancient DNA data, which would facilitate assessment of species diversity through time in this region.

In all study species except collared lemmings, we included samples from Wrangel Island. This island is well-known for holding the final refuge of mammoths until their extinction in the mid-late Holocene (Vartanyan, Garutt and Sher, 1993; Vartanyan *et al.*, 1995). In both the woolly rhinoceros and true lemmings, we observed distinct lineages on Wrangel Island compared with the mainland. Thus, isolation on Wrangel Island during warm periods may have contributed to higher rates of divergence in some Arctic species. Further whole genome samples from mainland Siberia for muskox will be required to test whether the lineage represented by the ancient sample was unique to Wrangel Island or was part of the wider Siberian diversity at the time. For true lemmings, mainland Siberian samples are required to further examine the phylogenetic relationships of the Siberian brown lemming. Additionally, ancient and modern DNA data from other species that have inhabited the island from the Late Pleistocene to today (e.g. *Dicrostonyx groenlandicus*), will be useful in exploring the role of isolation in the Wrangel Island refugium in species divergence.

More broadly, for collared and true lemmings, this thesis has focused primarily on the Late Pleistocene dynamics on the Eurasian side of the Northern Hemisphere, as large parts of this region remained ice-free during the LGM, with the exception of Scandinavia. However, to understand the evolutionary history of these genera in more detail, Late Pleistocene North

American samples are required. Furthermore, the genetic impact of range contractions on the North American counterparts of lemmings and muskox are currently largely unknown. It is likely that the fossil record for this region also harbours unique Pleistocene lineages that were lost during previous interglacials and the transition to the Holocene.

Implications for Arctic species under anthropogenic climate warming

While this thesis has focused on the impact of past climatic events, now, more than ever, it is imperative to look forward. The Arctic is presently warming at an unprecedented rate (Biskaborn *et al.*, 2019). Under the current scenarios of anthropogenic warming, Arctic species are at risk of a reduction in their remaining tundra habitat, due to the treeline moving northwards and towards areas of higher elevation, as well as changes in winter ice and snowfall. Arctic species are also facing competition from temperate species as the ranges of these species also move northward (Thulin and Tegelström, 2001; Gallant *et al.*, 2012; Elmhagen *et al.*, 2017). This may pose a threat to population numbers through competition for resources and expose species to interspecific hybridisation. Furthermore, lemmings are an important food source for Arctic predators and changes in their population cycling dynamics are already affecting Arctic fox populations (Angerbjörn *et al.*, 2013). A dramatic reduction in their numbers could have widespread consequences across the Arctic ecosystem. Of the species investigated here, anthropogenic warming will likely have the largest effect on muskox, which already persists with extremely low levels of genetic diversity and low effective population sizes in a fragmented distribution, thus having reduced potential to adapt to future environmental changes. Additionally, larger bodied species with longer generation times may not be able to adapt quickly enough, a factor that likely already contributed to the extinction of many megafauna across the globe, including the woolly rhinoceros. Thus, it is important to at least maintain the residual diversity of these species to give them the best chance of adapting to possible future threats. Future studies investigating the impacts of ongoing warming on Arctic species should incorporate palaeogenomic data indicating the loss of diversity and reduction of population size during interglacials such as the Eemian into species distribution modelling. This, along with changes in fossil distributions will aid in predicting not only how Arctic taxa may respond to ongoing climate warming, but also in understanding the wider changes in ecosystems that will occur as a result, as well as potentially identify focal areas or species for conservation action.

Conclusions

Overall, this thesis demonstrated that Arctic mammals had dynamic population histories and that the climatic fluctuations of the Late Pleistocene likely played a role in demographic changes across these species. Past climate, including sea level change and ice sheet dynamics influenced the demographic histories, population structure, speciation, and adaptation of Arctic taxa. Palaeogenomics has become an integral tool to assess the consequences of past climate change. This thesis showed that warming periods during the Late Pleistocene and transition to the Holocene had a marked effect on cold-adapted taxa, through reductions in population sizes and genetic diversity. The Eurasian collared lemming likely experienced a population

bottleneck in the Eemian interglacial, and subsequently diversified in the Last Glacial period. The woolly rhinoceros was also affected by warming, likely undergoing a rapid decline towards extinction associated with the Bølling-Allerød interstadial. The transition to the Holocene led to multiple founder events as muskox moved into the Canadian Arctic and Greenland, resulting in extremely low levels of genetic diversity in populations today. All species also showed a reduction in the number of mitochondrial lineages from the Pleistocene to the Holocene. Additionally, the LGM and formation of major ice sheets played a role in speciation and divergence in the Norwegian lemming and muskox. While this thesis has examined broad-scale patterns across Arctic species, future palaeogenomic studies might focus on specific climatic events, such as the Eemian interglacial and Bølling-Allerød interstadial, to characterise their impacts, as well as including multiple species to discern the effects of climate change across the Arctic ecosystem.

References

- Angerbjörn, A. *et al.* (2013) 'Carnivore conservation in practice: replicated management actions on a large spatial scale', *The Journal of applied ecology*, 50(1), pp. 59–67.
- Barlow, A. *et al.* (2018) 'Partial genomic survival of cave bears in living brown bears', *Nature Ecology & Evolution*, 2(10), pp. 1563–1570.
- Barnett, R. *et al.* (2009) 'Phylogeography of lions (*Panthera leo* ssp.) reveals three distinct taxa and a late Pleistocene reduction in genetic diversity', *Molecular ecology*, 18(8), pp. 1668–1677.
- Barr (1991) 'The commercial trade in muskox hides in the Northwest Territories 1860–1916', *Rangifer* [Preprint]. Available at: <https://septentrio.uit.no/index.php/rangifer/article/download/982/939>.
- Bennike, O. and Andreassen, C. (2005) 'Radiocarbon dating of musk-ox (*Ovibos moschatus*) remains from northeast Greenland', *The Polar record*, 41(219), pp. 305–310.
- Bergström, A. *et al.* (2022) 'Grey wolf genomic history reveals a dual ancestry of dogs', *Nature*, 607(7918), pp. 313–320.
- Biskaborn, B.K. *et al.* (2019) 'Permafrost is warming at a global scale', *Nature communications*, 10(1), p. 264.
- Bolger, A.M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic: a flexible trimmer for Illumina sequence data', *Bioinformatics*, 30(15), pp. 2114–2120.
- Brace, S. *et al.* (2012) 'Serial population extinctions in a small mammal indicate Late Pleistocene ecosystem instability', *Proceedings of the National Academy of Sciences*, 109(50), pp. 20532–20536.
- Butler, J. *et al.* (2008) 'ALLPATHS: de novo assembly of whole-genome shotgun microreads', *Genome research*, 18(5), pp. 810–820.
- Cahill, J.A. *et al.* (2018) 'Genomic evidence of widespread admixture from polar bears into brown bears during the last ice age', *Molecular biology and evolution*, 35(5), pp. 1120–1129.
- Campos, P.F. *et al.* (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*, 107(12), pp. 5675–5680.
- Carøe, C. *et al.* (2018) 'Single-tube library preparation for degraded DNA', *Methods in ecology and evolution / British Ecological Society*. Edited by S. Johnston, 9(2), pp. 410–419.
- Cingolani, P. *et al.* (2012) 'A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3', *Fly*, 6(2), pp. 80–92.
- Clark, P.U. *et al.* (2009) 'The Last Glacial Maximum', *Science*, 325(5941), pp. 710–714.
- Crump, S.E. *et al.* (2021) 'Ancient plant DNA reveals High Arctic greening during the Last Interglacial', *Proceedings of the National Academy of Sciences of the United States of America*, 118(13). doi:10.1073/pnas.2019069118.
- Cuyler, C. *et al.* (2020) 'Muskox status, recent variation, and uncertain future', *Ambio*, 49(3), pp. 805–819.
- Dalén, L. *et al.* (2005) 'Population history and genetic structure of a circumpolar species: the arctic fox', *Biological journal of the Linnean Society. Linnean Society of London*, 84(1), pp. 79–89.
- Davydov, E.V. *et al.* (2010) 'Identifying a high fraction of the human genome to be under selective constraint using GERP++', *PLoS computational biology*, 6(12), p. e1001025.
- Dehasque, M. *et al.* (2021) 'Combining Bayesian age models and genetics to investigate population dynamics and extinction of the last mammoths in northern Siberia', *Quaternary science reviews*, 259, p. 106913.
- Dehasque, M. *et al.* (2022) 'Development and Optimization of a Silica Column-Based Extraction Protocol for Ancient DNA', *Genes*, 13(4). doi:10.3390/genes13040687.
- Drummond, A.J. and Rambaut, A. (2007) 'BEAST: Bayesian evolutionary analysis by sampling trees', *BMC evolutionary biology*, 7(1), p. 214.
- Durand, E.Y. *et al.* (2011) 'Testing for ancient admixture between closely related populations', *Molecular biology and evolution*, 28(8), pp. 2239–2252.
- Dussex, N. *et al.* (2020) 'Moose genomes reveal past glacial demography and the origin of modern lineages', *BMC genomics*, 21(1), p. 854.
- Dussex, N. *et al.* (2021) 'Population genomics of the critically endangered kākāpō', *Cell Genomics*, 1(1), p. 100002.
- Eden, E. *et al.* (2009) 'GORilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists', *BMC bioinformatics*, 10, p. 48.
- Edwards, M.E. (2020) 'The maturing relationship between Quaternary paleoecology and ancient sedimentary DNA', *Quaternary Research*, 96, pp. 39–47.

- Elmhagen, B. *et al.* (2017) 'Homage to Hersteinsson and Macdonald: climate warming and resource subsidies cause red fox range expansion and Arctic fox decline', *Polar research*, 36(sup1), p. 3.
- Ersmark, E. *et al.* (2019) 'Genetic turnovers and northern survival during the last glacial maximum in European brown bears', *Ecology and evolution*, 9(10), pp. 5891–5905.
- Excoffier, L. *et al.* (2021) 'fastsimcoal2: demographic inference under complex evolutionary scenarios', *Bioinformatics* [Preprint]. doi:10.1093/bioinformatics/btab468.
- Fages, A. *et al.* (2019) 'Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series', *Cell*, 177(6), pp. 1419–1435.e31.
- Fedorov, V. *et al.* (1999) 'Phylogeographic structure and mitochondrial DNA variation in true lemmings (*Lemmus*) from the Eurasian Arctic', *Biological journal of the Linnean Society. Linnean Society of London*, 66(3), pp. 357–371.
- Fedorov, V.B. *et al.* (2020) 'Impact of past climate warming on genomic diversity and demographic history of collared lemmings across the Eurasian Arctic', *Proceedings of the National Academy of Sciences of the United States of America*, 117(6), pp. 3026–3033.
- Fedorov, V. and Stenseth, N.C. (2001) 'Glacial survival of the Norwegian lemming (*Lemmus lemmus*) in Scandinavia: inference from mitochondrial DNA variation', *Proceedings. Biological sciences / The Royal Society*, 268(1469), pp. 809–814.
- Feng, S. *et al.* (2020) 'Dense sampling of bird diversity increases power of comparative genomics', *Nature*, 587(7833), pp. 252–257.
- Gallant, D. *et al.* (2012) 'Arctic fox versus red fox in the warming Arctic: four decades of den surveys in north Yukon', *Polar biology*, 35(9), pp. 1421–1431.
- Gelabert, P. *et al.* (2021) 'Genome-scale sequencing and analysis of human, wolf, and bison DNA from 25,000-year-old sediment', *Current biology: CB*, 31(16), pp. 3564–3574.e9.
- Graham, R.W. *et al.* (2016) 'Timing and causes of mid-Holocene mammoth extinction on St. Paul Island, Alaska', *Proceedings of the National Academy of Sciences*, 113(33), pp. 9310–9314.
- Green, R.E. *et al.* (2010) 'A draft sequence of the Neandertal genome', *Science*, 328(5979), pp. 710–722.
- Hahn, C., Bachmann, L. and Chevreux, B. (2013) 'Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach', *Nucleic acids research*, 41(13), p. e129.
- Hansen, C.C.R. *et al.* (2018) 'The Muskox Lost a Substantial Part of Its Genetic Diversity on Its Long Road to Greenland', *Current biology: CB*, 28(24), pp. 4022–4028.e5.
- Harington, C.R. (1980) 'Radiocarbon dates on some Quaternary mammals and artifacts from northern North America', *Arctic*, 33(4), pp. 815–832.
- Hedrick, P.W. (2001) 'Conservation genetics: where are we now?', *Trends in ecology & evolution*, 16(11), pp. 629–636.
- Hewitt, G. (2000) 'The genetic legacy of the Quaternary ice ages', *Nature*, 405(6789), pp. 907–913.
- Hewitt, G.M. (1996) 'Some genetic consequences of ice ages, and their role in divergence and speciation', *Biological journal of the Linnean Society. Linnean Society of London*, 58(3), pp. 247–276.
- Hirata, D. *et al.* (2017) 'Paternal phylogeographic structure of the brown bear (*Ursus arctos*) in northeastern Asia and the effect of male-mediated gene flow to insular populations', *Zoological letters*, 3, p. 21.
- Hofreiter, M. *et al.* (2004) 'Lack of phylogeography in European mammals before the last glaciation', *Proceedings of the National Academy of Sciences of the United States of America*, 101(35), pp. 12963–12968.
- Huber, C.D., Kim, B.Y. and Lohmueller, K.E. (2020) 'Population genetic models of GERP scores suggest pervasive turnover of constrained sites across mammalian evolution', *PLoS genetics*, 16(5), p. e1008827.
- Jia, P. *et al.* (2019) 'Four Novel SNPs of MYO1A Gene Associated with Heat-Tolerance in Chinese Cattle', *Animals : an open access journal from MDPI*, 9(11). doi:10.3390/ani9110964.
- Kearse, M. *et al.* (2012) 'Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data', *Bioinformatics*, 28(12), pp. 1647–1649.
- Kirillova, I.V. *et al.* (2015) 'An ancient bison from the mouth of the Rauchua River (Chukotka, Russia)', *Quaternary Research*, 84(2), pp. 232–245.
- Klüttsch, C.F.C. *et al.* (2017) 'Evolutionary reconstruction supports the presence of a Pleistocene Arctic refugium for a large mammal species', *Journal of biogeography*, 44(12), pp. 2729–2739.
- Knuth, E. (1967) *Archaeology of the musk-ox way*. Sorbonne: École pratique des hautes études.
- Korneliussen, T.S., Albrechtsen, A. and Nielsen, R. (2014) 'ANGSD: Analysis of Next Generation Sequencing Data', *BMC bioinformatics*, 15, p. 356.
- Kutschera, V.E. *et al.* (2022) 'GenErode: a bioinformatics pipeline to investigate genome erosion in endangered

- and extinct species', *BMC bioinformatics*, 23(1), p. 228.
- Lagerholm, V.K. *et al.* (2014) 'On the origin of the Norwegian lemming', *Molecular ecology*, 23(8), pp. 2060–2071.
- Larsen, E. *et al.* (1987) 'Cave stratigraphy in western Norway; multiple Weichselian glaciations and interstadial vertebrate fauna', *Boreas*, 16(3), pp. 267–292.
- Larsson, P. *et al.* (2019) 'Consequences of past climate change and recent human persecution on mitogenomic diversity in the arctic fox', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 374(1788), p. 20190212.
- Lent, P.C. (1999) *Muskoxen and Their Hunters: A History*. University of Oklahoma Press.
- Li, H. and Durbin, R. (2011) 'Inference of human population history from individual whole-genome sequences', *Nature*, 475(7357), pp. 493–496.
- Li, R. *et al.* (2010) 'De novo assembly of human genomes with massively parallel short read sequencing', *Genome research*, 20(2), pp. 265–272.
- Lister, A.M. and Sher, A.V. (2001) 'The origin and evolution of the woolly mammoth', *Science*, 294(5544), pp. 1094–1097.
- Lister, A.M. and Stuart, A.J. (2008) 'The impact of climate change on large mammal distribution and extinction: Evidence from the last glacial/interglacial transition', *Comptes rendus: Geoscience*, 340(9), pp. 615–620.
- Liu, S. *et al.* (2014) 'Population Genomics Reveal Recent Speciation and Rapid Evolutionary Adaptation in Polar Bears', *Cell*, 157(4), pp. 785–794.
- Liu, S. *et al.* (2021) 'Ancient and modern genomes unravel the evolutionary history of the rhinoceros family', *Cell*, 184(19), pp. 4874–4885.
- Loog, L. *et al.* (2020) 'Ancient DNA suggests modern wolves trace their origin to a Late Pleistocene expansion from Beringia', *Molecular ecology*, 29(9), pp. 1596–1610.
- Lorenzen, E.D. *et al.* (2011) 'Species-specific responses of Late Quaternary megafauna to climate and humans', *Nature*, 479(7373), pp. 359–364.
- Lynch, V.J. *et al.* (2015) 'Elephantid Genomes Reveal the Molecular Bases of Woolly Mammoth Adaptations to the Arctic', *Cell reports*, 12(2), pp. 217–228.
- Markova, A.K. *et al.* (1995) 'Late Pleistocene distribution and diversity of mammals in Northern Eurasia (PALEOFAUNA database)'. Available at: <https://elibrary.ru/item.asp?id=21313678>.
- Martin, S.H. and Van Belleghem, S.M. (2017) 'Exploring Evolutionary Relationships Across the Genome Using Topology Weighting', *Genetics*, 206(1), pp. 429–438.
- Mather, N., Traves, S.M. and Ho, S.Y.W. (2020) 'A practical introduction to sequentially Markovian coalescent methods for estimating demographic history from genomic data', *Ecology and evolution*, 10(1), pp. 579–589.
- Mathur, S. and DeWoody, J.A. (2021) 'Genetic load has potential in large populations but is realized in small inbred populations', *Evolutionary applications*, 14(6), pp. 1540–1557.
- Mays, H.L. *et al.* (2018) 'Genomic Analysis of Demographic History and Ecological Niche Modeling in the Endangered Sumatran Rhinoceros *Dicerorhinus sumatrensis*', *Current biology: CB*, 28(1), pp. 70–76.e4.
- Mazet, O. *et al.* (2016) 'On the importance of being structured: instantaneous coalescence rates and human evolution--lessons for ancestral population size inference?', *Heredity*, 116(4), pp. 362–371.
- Meiri, M. *et al.* (2020) 'Population dynamics and range shifts of moose (*Alces alces*) during the Late Quaternary', *Journal of biogeography*, 38, p. 29.
- Meisner, J. and Albrechtsen, A. (2018) 'Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data', *Genetics*, 210(2), pp. 719–731.
- Meyer, M. and Kircher, M. (2010) 'Illumina sequencing library preparation for highly multiplexed target capture and sequencing', *Cold Spring Harbor protocols*, 2010(6), p. db. prot5448.
- Mi, H. *et al.* (2013) 'Large-scale gene function analysis with the panther classification system', *Nature protocols*, 8(8), pp. 1551–1566.
- Mitchell, K.J. and Rawlence, N.J. (2021) 'Examining Natural History through the Lens of Palaeogenomics', *Trends in ecology & evolution* [Preprint]. doi:10.1016/j.tree.2020.10.005.
- Morin, P.A. *et al.* (2021) 'Reference genome and demographic history of the most endangered marine mammal, the vaquita', *Molecular ecology resources*, 21(4), pp. 1008–1020.
- Nota, K. *et al.* (2022) 'Norway spruce postglacial recolonization of Fennoscandia', *Nature communications*, 13(1), p. 1333.
- Nyström, V. *et al.* (2010) 'Temporal genetic change in the last remaining population of woolly mammoth', *Proceedings. Biological sciences / The Royal Society*, 277(1692), pp. 2331–2337.

- Orlando, L. *et al.* (2013) 'Recalibrating equus evolution using the genome sequence of an early Middle Pleistocene horse', *Nature*, 499(7456), pp. 74–78.
- Pacz, S. *et al.* (2022) 'Reference genomes for conservation', *Science*, 377(6604), pp. 364–366.
- Palkopoulou, E. *et al.* (2013) 'Holarctic genetic structure and range dynamics in the woolly mammoth', *Proceedings of the Royal Society B: Biological Sciences*, 280(1770), pp. 20131910–20131910.
- Palkopoulou, E. *et al.* (2015) 'Complete genomes reveal signatures of demographic and genetic declines in the woolly mammoth', *Current biology: CB*, 25(10), pp. 1395–1400.
- Palkopoulou, E. *et al.* (2016) 'Synchronous genetic turnovers across Western Eurasia in Late Pleistocene collared lemmings', *Global change biology*, 22(5), pp. 1710–1721.
- Patterson, N. *et al.* (2012) 'Ancient admixture in human history', *Genetics*, 192(3), pp. 1065–1093.
- Pečnerová, P. *et al.* (2017) 'Mitogenome evolution in the last surviving woolly mammoth population reveals neutral and functional consequences of small population size', *Evolution Letters*, 1(6), pp. 292–303.
- Pečnerová, P. (2018) *Genomic analysis of the process leading up to the extinction of the woolly mammoth*. Stockholm University.
- Pečnerová, P. *et al.* (2021) 'High genetic diversity and low differentiation reflect the ecological versatility of the African leopard', *Current biology: CB*, 31(9), pp. 1862–1871.e5.
- Petkova, D., Novembre, J. and Stephens, M. (2016) 'Visualizing spatial population structure with estimated effective migration surfaces', *Nature genetics*, 48(1), pp. 94–100.
- Pickrell, J.K. and Pritchard, J.K. (2012) 'Inference of population splits and mixtures from genome-wide allele frequency data', *PLoS genetics*, 8(11), p. e1002967.
- Prost, S. *et al.* (2010) 'Influence of climate warming on arctic mammals? new insights from ancient DNA studies of the collared lemming *Dicrostonyx torquatus*', *PloS one*, 5(5), p. e10447.
- Purcell, S. *et al.* (2007) 'PLINK: a tool set for whole-genome association and population-based linkage analyses', *American journal of human genetics*, 81(3), pp. 559–575.
- Ramsey, C.B. (2009) 'Bayesian Analysis of Radiocarbon Dates', *Radiocarbon*, 51(1), pp. 337–360.
- Renaud, G. *et al.* (2019) 'Joint Estimates of Heterozygosity and Runs of Homozygosity for Modern and Ancient Samples', *Genetics*, 212(3), pp. 587–614.
- Rey-Iglesia, A. *et al.* (2021) 'Late Pleistocene paleoecology and phylogeography of woolly rhinoceroses', *Quaternary science reviews*, 263, p. 106993.
- Rhie, A. *et al.* (2021) 'Towards complete and error-free genome assemblies of all vertebrate species', *Nature*, 592(7856), pp. 737–746.
- Rogers, R.L. and Slatkin, M. (2017) 'Excess of genomic defects in a woolly mammoth on Wrangel island', *PLoS genetics*. Edited by G.S. Barsh, 13(3), p. e1006601.
- Seppy, M., Manni, M. and Zdobnov, E.M. (2019) 'BUSCO: Assessing Genome Assembly and Annotation Completeness', in Kollmar, M. (ed.) *Gene Prediction: Methods and Protocols*. New York, NY: Springer New York, pp. 227–245.
- von Seth, J. *et al.* (2021) 'Genomic insights into the conservation status of the world's last remaining Sumatran rhinoceros populations', *Nature communications*, 12(1), p. 2393.
- Sikora, M. *et al.* (2019) 'The population history of northeastern Siberia since the Pleistocene', *Nature*, 570(7760), pp. 182–188.
- Simpson, J.T. *et al.* (2009) 'ABYSS: a parallel assembler for short read sequence data', *Genome research*, 19(6), pp. 1117–1123.
- Sinding, M.-H.S. *et al.* (2018) 'Population genomics of grey wolves and wolf-like canids in North America', *PLoS genetics*, 14(11), p. e1007745.
- Skoglund, P. *et al.* (2015) 'Genetic evidence for two founding populations of the Americas', *Nature*, 525(7567), pp. 104–108.
- Skotte, L., Korneliussen, T.S. and Albrechtsen, A. (2013) 'Estimating individual admixture proportions from next generation sequencing data', *Genetics*, 195(3), pp. 693–702.
- Spyrou, M.A. *et al.* (2019) 'Ancient pathogen genomics as an emerging tool for infectious disease research', *Nature reviews: Genetics*, 20(6), pp. 323–340.
- Stewart, J.R. *et al.* (2010) 'Refugia revisited: individualistic responses of species in space and time', *Proceedings. Biological sciences / The Royal Society*, 277(1682), pp. 661–671.
- Stokes, C.R. (2017) 'Deglaciation of the Laurentide Ice Sheet from the Last Glacial Maximum', *Cuadernos de investigación geográfica*, 43(2), p. 0.
- Stuart, A.J. and Lister, A.M. (2007) 'Patterns of Late Quaternary megafaunal extinctions in Europe and northern

- Asia', *Courier-Forschungsinstitut Senckenberg*, 259, pp. 287–297.
- Stuart, A.J. and Lister, A.M. (2012) 'Extinction chronology of the woolly rhinoceros *Coelodonta antiquitatis* in the context of late Quaternary megafaunal extinctions in northern Eurasia', *Quaternary Science Reviews*. Pergamon, pp. 1–17. doi:10.1016/j.quascirev.2012.06.007.
- Svensson, A. *et al.* (2006) 'The Greenland Ice Core Chronology 2005, 15–42 ka. Part 2: comparison to other records', *Quaternary science reviews*, 25(23–24), pp. 3258–3267.
- Thulin, C.-G. and Tegelström, H. (2001) 'High mtDNA haplotype diversity among introduced Swedish brown hares *Lepus europaeus*', *Acta theriologica*, 46(4), pp. 375–384.
- To, T.-H. *et al.* (2016) 'Fast Dating Using Least-Squares Criteria and Algorithms', *Systematic biology*, 65(1), pp. 82–97.
- Tunstall, T. *et al.* (2018) 'Evaluating recovery potential of the northern white rhinoceros from cryopreserved somatic cells', *Genome research*, 28(6), pp. 780–788.
- van der Valk, T. *et al.* (2019) 'Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in Eastern Gorillas', *Current biology: CB*, 29(1), pp. 165–170.e6.
- van der Valk, T. *et al.* (2021) 'Million-year-old DNA sheds light on the genomic history of mammoths', *Nature*, 591(7849), pp. 265–269.
- Vartanyan, S.L. *et al.* (1995) 'Radiocarbon Dating Evidence for Mammoths on Wrangel Island, Arctic Ocean, Until 2000 BC1', *Radiocarbon*, 37(1), pp. 1–6.
- Vartanyan, S.L., Garutt, V.E. and Sher, A.V. (1993) 'Holocene dwarf mammoths from Wrangel Island in the Siberian Arctic', *Nature*, 362(6418), pp. 337–340.
- Vershinina, A.O. *et al.* (2021) 'Ancient horse genomes reveal the timing and extent of dispersals across the Bering Land Bridge', *Molecular ecology*, 30(23), pp. 6144–6161.
- Voldstad, L.H. *et al.* (2020) 'A complete Holocene lake sediment ancient DNA record reveals long-standing high Arctic plant diversity hotspot in northern Svalbard', *Quaternary science reviews*, 234, p. 106207.
- Waltari, E. and Cook, J.A. (2005) 'Hares on ice: phylogeography and historical demographics of *Lepus arcticus*, *L. othus*, and *L. timidus* (Mammalia: Lagomorpha)', *Molecular ecology*, 14(10), pp. 3005–3016.
- Willerslev, E. *et al.* (2003) 'Diverse plant and animal genetic records from Holocene and Pleistocene sediments', *Science*, 300(5620), pp. 791–795.
- Yang, D.Y. *et al.* (1998) 'Technical Note: Improved DNA Extraction From Ancient Bones Using Silica-Based Spin Columns', 543(December 1997), pp. 539–543.

Acknowledgements

They say it takes a village to raise a child, and a thesis is no different. Firstly, I would like to acknowledge Nat, Clare, Denise, Susie, and Anu for being the inspiration and drive to undertake this PhD experience in Sweden. If it wasn't for you girls, I wouldn't have had the confidence to move to the other side of the planet for science. Thank you for all the zoom dates, whatsapp updates, enthusiasm for ice cream, being the best role models, as well as your life and academic advice. I miss you all very much and can't wait for the big post-theses adventure!

Of course, no PhD could be undertaken without a supervisor, and an excellent one at that. Love, thank you for accepting me into this PhD and for always supporting and pushing me to strive for my best. I will always appreciate your invaluable advice, throwing me side projects, your enthusiasm for all things lemmings, your many stories over a beer, and being serenaded with Lucia songs by a whole choir. I've learnt more than I could have ever thought in this PhD and had a lot of fun doing it.

To the postdocs who helped drive the projects in my thesis - Nic, David D, and Patricia - it has been an honour to work with you over these years. Nic, thanks for chatting with me about the lab before I started, and especially for stating "if you like board games, it's a fun place to be", and for being someone I could talk to about NZ related things. David, I know I'll never be able to say rhinoceros properly when you're around (or ever). Thanks for all your help, patience, bioinformatics advice, tiktok quoting, and for making the group environment what it is. Patricia, thanks for sharing your muskox enthusiasm and projects with me, and for being an inspiration. A special thank you to the Bolin Centre for Climate Research for funding several muskox side projects, one of which end up in the thesis.

A big thank you to Verena, Tom, and Per for additional bioinformatic support and advice. Verena, you fielded my many many questions about the pipeline, especially when I had to break it to make it do what I needed. Tom, thank you for always having an answer to a question and for "just writing a little python script" when needed. Thank you also to Anders for co-supervising, and for getting me to sequence some "Plague guys" in a very short time frame. One day we'll get round to that human project that was supposed to be in my thesis. Finally, thank you to Martin, Lars, and Barbara for your valuable input as part of the follow up committee.

I'd like to thank all the collaborators involved in getting the samples that I've worked with during this PhD. And, a big thank you to Denise, Nat, Johanna, Jorge, Marianne, Mário, and Clare for your comments and proofreading of this thesis summary, and to Erik and Johanna for proofreading the Swedish abstract.

To Marianne, Johannes, and Johanna, my fellow Lover PhD buddies over the past four years - I cannot have imagined a better bunch of people to get through a PhD with. Thank you for your support, kind words, hilarious phrases, bad puns, ridiculous amounts of beer, and multitude of

shenanigans. We've laughed and cried together, and I'm proud of us. I can't wait to see what you all get up to post-PhD.

To David D, Marianne, Johannes, Johanna, Mário, Ioana, Zoe, Nora, Isa, Camilo, Tom, Ernst, Nic, Anna, Petter, Reyhan, Erik, Maja, TJ, and Ricardo - thanks for your support, sarcasm, office discussions, and laughter over the past years, and for keeping me somewhat sane (okay sane is a stretch) during the final months of this PhD. Also to the CPGers and the NRM crew past and present - thank you all for the over 570 recorded (and counting!) games of 'Bang!' and all the Bang-related lingo ("elegant" can never be used in any other way). Also for the memorable lunch and fika discussions, New Brew Thursdays (thanks to Johannes for always having the fridge well-stocked), frisbee, caving (I hope your stress levels have decreased, Ioana), and Nut Island adventures (shoutout to Marianne for getting us back to Stockholm at 2am unexpectedly on the last trip). Thanks also to the wider Zoology PhD students for your support, Friday pubs/cocktails, and Tovetorp midsummers the past couple of years.

To David M(eep!), Maarit and Jorge (oulala!), Philip and Jovanka, Fanny and Tea, Ingo, Ciaran, and Natalia - thank you for being my go-to support crew. We've enjoyed many a weekend wander, dinner, brunch, forest foraging, wine, beer, or whiskey together. You guys are all awesome, and I've really appreciated all the advice. The Fanny island, Lapland, Viborg, and Tarago trips were certainly highlights during this time.

And finally, a big thank you to Eric, Julie, Regan, and Hollie, and the rest of my family and friends in New Zealand. Your support from afar is very much appreciated. I've missed you all very much and will come and visit again soon.

Svensk sammanfattning

Klimatförändringarna under Övre pleistocen har troligen haft en stor inverkan på de arktiska arternas evolutionära historia. Paleogenomik är ett användbart verktyg för att belysa hur tidigare populationer påverkades av dessa klimatförändringar och den sammankopplade dynamiken i istäcken och förändringen av havsnivån. Här har jag använt moderna och förhistoriska DNA-data från fyra arktiska däggdjur för att undersöka effekterna av det senpleistocena klimatet på deras evolutionära historia, från populationsdynamik och demografi till artbildning och genflöde, anpassning och genomisk erosion. I *artikel I* visade jag med hjälp av gamla mitokondriella arvsmassor från hela deras senpleistocena utbredningsområde, att den eurasiska halsbandslämmeln (*Dicrostonyx torquatus*) hade en dynamisk senpleistocen populationsstruktur i Europa. Dessutom ledde den Eemiska interglacialen sannolikt till en flaskhals hos halsbandslämlar, varefter arten diversifierades under den senaste istiden. Nukleär genomdata från en modern individ i nordöstra Sibirien tyder på populationsstabilitet i nordöstra Sibirien under holocen. I *artikel II* sekvenserade jag arvsmassan hos en ~18 500 år gammal ullhårig noshörning (*Coelodonta antiquitatis*) och använde detta i kombination med mitokondriella data för att utforska artens demografiska historia. Det fanns liten geografisk strukturering i den nordostsibiriska populationen och stabilitet i deras effektiva populationsstorlek strax före utrotningen, vilket kan tyda på en efterföljande snabb nedgång mot utrotning som troligen var förknippad med Bølling-Allerød-interstadialen. Dessutom fann jag att denna art hade mutationer i genen TRPA1, som är inblandad i temperaturkänslighet. I en tredje studie (*artikel III*) använde jag hela arvsmassor från moderna och gamla äkta lämlar för att fastställa att den norska lämmeln har en av de yngsta artbildningstiderna (~37-34 ka BP) bland däggdjur. Norska lämlar har mutationer i gener som är involverade i pälsfärg, färguppfattning, fettransport och reproduktion, och har sannolikt utvecklat sin unika färgteckning som ett resultat av isolering efter rekoloniseringen av Fennoskandien. Slutligen undersökte vi konsekvenserna av en långsiktigt liten effektiv populationsstorlek hos myskoxen (*Ovibos moschatus*) med hjälp av 107 moderna arvsmassor och en 21 000 år gammal Sibiriskt arvsmassa (*artikel IV*). Myskoxen överlevde uppvärmningen i slutet av senpleistocen, men de successiva grundareffekterna som inträffade under koloniseringen av Norra Kanada och Grönland, minskade den genetiska mångfalden till några av de lägsta värden som observerats hos däggdjur. Resultaten tyder dock på att det faktum att populationen varit liten under lång tid troligen ledde till att starkt skadliga alleler rensades bort hos myskoxen, vilket i sin tur gjorde att de kunde överleva med begränsade tecken på inavel. Ur teknisk synvinkel presenteras i denna avhandling fyra *de-novo* referensgenomkomplex och de första kompletta nukleära arvsmassorna för de undersökta arterna. Sammantaget visar resultaten i denna avhandling att klimatförändringarna, särskilt Eem-istiden och Bølling-Allerød-istiden, tillsammans med förändringar i havsnivå samt bildandet och tillbakadragandet av istäcken under den senaste istiden, har påverkat dessa fyra arktiska däggdjurs evolutionära historia.