Genomic analysis of the process leading up to the extinction of the woolly mammoth

Patricia Pečnerová

Academic dissertation for the Degree of Doctor of Philosophy in Systematic Zoology at Stockholm University to be publicly defended on Tuesday 13 February 2018 at 10.00 in Vivi Täckholmssalen, NPQ-huset, Svante Arrhenius väg 20.

Abstract
Species worldwide are subject to contractions in both abundance and geographical range, and their persistence in a changing environment may thus depend on the ability to survive in small and fragmented populations. Despite the urgent need to understand how extinction works, our knowledge of pre-extinction genetic processes is limited because techniques allowing population and conservation genomics to be studied in wild threatened populations have become available only recently. In this thesis, I used the last surviving population of the woolly mammoth (Mammuthus primigenius) as a model for studying pre-extinction population dynamics. I used ancient DNA as a tool to study microevolutionary processes in real time, analysing genetic changes in response to environmental shifts at the end of the last Ice Age and exploring impacts of genetic drift and inbreeding as woolly mammoths became isolated on Wrangel Island and survived for 6000 years at small population size. Using mitochondrial genomes, I found evidence of a founder effect that decreased the maternal diversity to a single lineage at the time when mammoths became trapped on Wrangel Island (~10,500 years ago). Moreover, a two- to three-fold higher mitochondrial mutation rate in Holocene and a fixed, potentially detrimental mutation in the ATP6 gene encoding for one of the key enzymes of the oxidative phosphorylation pathway, is consistent with the hypothesis that selection is less effective in removing deleterious mutations in small populations. A loss of diversity was also observed in an immunity gene that belongs to the major histocompatibility complex (MHC), even though the MHC is considered to be under balancing selection. Low-coverage genomic data was analysed in order to estimate endogenous DNA content and molecular sex of the mammoth samples. The observation of a male bias (69%) in the sex ratio led to the conclusion that male mammoths were more likely to die in a way that ensured good preservation. Another potential way of getting information about life history strategies of extinct species, which was explored here, is by measuring testosterone levels in mammoth hair shafts in connection with molecular sex inference. Finally, given that previous estimates have suggested a very small Holocene effective population size on Wrangel Island and thus that the population may have been too small to avoid genome erosion, four mammoths were sequenced to a high coverage in order to look for genomic consequences of small population size. When compared to mammoths from the Pleistocene mainland population, Wrangel Island mammoths had lower levels of genome-wide diversity and had a higher proportion of their genomes allocated in runs of homozygosity, which are large fragments completely depleted of diversity. Importantly, genome erosion appears to have accelerated in the last ten generations before the extinction, resulting in the last known woolly mammoth having almost 40% of its genome without any genetic diversity. Overall, these results highlight how genetic drift and inbreeding triggered genomic deterioration in the last surviving woolly mammoth population. Although Wrangel Island was a refugium, where mammoths survived for thousands of years after the last Ice Age, and the causal factors of the final extinction are not yet clear, isolation and small population size without any possibility of new gene flow may have contributed to reduced fitness, and thus to extinction.

Keywords: woolly mammoth, Mammuthus primigenius, extinction, Wrangel Island, ancient DNA, palaeogenetics, population genetics, genomics, genomic erosion, genetic drift, inbreeding, population size.

Stockholm 2018
http://urn.kb.se/resolve?urn=urn:nbn:se:diva-149655


Department of Zoology
Stockholm University, 106 91 Stockholm
GENOMIC ANALYSIS OF THE PROCESS LEADING UP TO THE EXTINCTION OF THE WOOLLY MAMMOTH

Patícia Pečnerová
Genomic analysis of the process leading up to the extinction of the woolly mammoth

Patricia Pečnerová
"Far away to the North-East in the ice-infested waters of the Chukchi Sea there is a big Soviet island. And though nature there is harsh and wild, almost 10 months a fierce winter rules there, and the sea bathing the island is cold seething and even in summer is covered by floating ice. This island is truly amazing. Everybody, who is lucky to work or stay there, will love it with heart and soul and will never forget this Unexplored Land."

- Aref Mineev, founder of the first wildlife reserve on Wrangel Island
List of papers

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


* These authors contributed equally to this work.

The articles are reprinted with permission from the respective publishers.
I am also a co-author of the following articles, which are not included in this thesis:


Candidate’s Contribution

Candidate contributions to thesis articles*

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceived the study</td>
<td>Significant</td>
<td>Significant</td>
<td>Substantial</td>
<td>Minor</td>
<td>Substantial</td>
</tr>
<tr>
<td>Designed the study</td>
<td>Substantial</td>
<td>Significant</td>
<td>Substantial</td>
<td>Minor</td>
<td>Substantial</td>
</tr>
<tr>
<td>Collected the data</td>
<td>Substantial</td>
<td>Substantial</td>
<td>Substantial</td>
<td>Significant</td>
<td>Substantial</td>
</tr>
<tr>
<td>Analysed the data</td>
<td>Substantial</td>
<td>Substantial</td>
<td>Substantial</td>
<td>Significant</td>
<td>Significant</td>
</tr>
<tr>
<td>Manuscript preparation</td>
<td>Substantial</td>
<td>Substantial</td>
<td>Substantial</td>
<td>Significant</td>
<td>Substantial</td>
</tr>
</tbody>
</table>

* 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.
# Table of Contents

- **Introduction** ................................................................................................... 15
- **Extinction factors** ................................................................................................. 16
- **The genetics of extinction** .............................................................................................. 16
- **Conservation palaeogenomics** .................................................................................... 18
- **Study system** ............................................................................................................ 18
- **Objectives** ............................................................................................................ 22

## Methods ....................................................................................................................... 23
- **Type of material** ...................................................................................................... 23
- **Dating** .................................................................................................................. 23
- **Ancient DNA** ........................................................................................................ 23
- **DNA extractions** .................................................................................................. 26
- **PCR** ...................................................................................................................... 26
- **High-Throuhgout Sequencing** .............................................................................. 27
- **Bioinformatics** ..................................................................................................... 28

## Results & Discussion ............................................................................................... 30
- **Demographic history of the Wrangel Island mammoths** ......................................... 30
- **Functional consequences of the bottleneck** ............................................................. 32
- **Behaviour and social structure** ............................................................................... 34
- **Genome erosion** .................................................................................................... 35

## Concluding remarks ................................................................................................. 38

## References ................................................................................................................... 40

- **Svensk sammanfattning** .......................................................................................... 49
- **Zhrnutie v slovenčine** ............................................................................................. 51

## Acknowledgements .................................................................................................... 53
Introduction

In order to understand and predict species’ responses to climatic and environmental changes, researchers have been looking for answers in the past 50,000 years of the fossil record. One way of tracking the impact of past climatic shifts and spreading human populations on species’ abundances is by generating extensive series of radiocarbon dates (Stuart et al. 2002; Cooper et al. 2015), which serve as an indication of presence or absence of taxa. Furthermore, the use of ancient DNA (aDNA) has revealed that many previously undetected, genetically distinct (mitochondrial) lineages went extinct in the Late Pleistocene (Barnes et al. 2002; Leonard et al. 2007; Palkopoulou et al. 2016).

Considering that the Late Quaternary population turnovers and extinctions involved about 90 genera of megafauna worldwide, a number of causes have been proposed to explain this complex phenomenon (Koch & Barnosky 2006; Stuart 2015). While the hyperdisease and extraterrestrial impact (MacPhee 1997; Firestone et al. 2007) hypotheses are considered as unlikely explanations for the Late Quaternary extinctions (Lister & Stuart 2008), the debate has generally revolved around the respective roles of climatic change and human hunting. This debate originated with the ‘overkill’ hypothesis (Martin 1967), and support has been piling up on both sides (Sandom et al. 2014; Cooper et al. 2015). As a result, many researchers have turned to a more comprehensive model, which combines the long-term process of range contractions during times of unfavourable conditions and the proximate extinction factors within terminal refugial populations (Barnosky et al. 2004). Refugia are considered as ‘core’ areas, where species retreated during episodes of unfavourable conditions (Lister & Stuart 2008). However, further range contractions resulted in smaller refugial populations, which were more sensitive to various stochastic factors. Understanding what was happening within the terminal refugia is thus critical for determining the final causes of extinction (Lister & Stuart 2008).
Extinction factors

The processes operating on a longer time scale and resulting in reduced population size, such as changing climatic conditions and receding habitats, are considered deterministic factors (also known as the declining-population paradigm). When a population becomes small, it is also exposed to stochastic factors, which can be environmental, catastrophic, demographic, and genetic (also known as the small-population paradigm; Frankel & Soulé 1981; Shaffer 1981; Lande 1993; Caughley 1994).

Extinction is usually a result of the interplay between deterministic and stochastic factors (Hedrick et al. 1996), and this might also have been the case in the Late Quaternary megafauna extinctions. Changing environmental conditions at the end of Pleistocene pushed species adapted to steppe-tundra habitat to northern refugia, decreasing their population sizes and making them susceptible to stochastic factors. Small and isolated populations were not only more vulnerable to human hunting, but also to fluctuations in the biotic environment (e.g. competition, predation, and diseases), natural catastrophes (e.g. floods or droughts), and demographic stochasticity in the survival and reproductive success (Shaffer 1981; Lande 1993). Furthermore, over the last four decades, genetic factors have been recognized as a threat to viability and have been hypothesized to play a considerable role in driving populations to extinction (Lande 1988; Frankham 2005).

If various extinction factors operate in parallel, their impacts are reinforced, which leads to a further reduction in the population size. This positive feedback loop is known as the ‘extinction vortex’ (Gilpin & Soulé 1986). According to the extinction vortex model, time to extinction scales logarithmically to population size, meaning that populations decline faster closer to extinction (Fagan & Holmes 2006).

The genetics of extinction

At small population sizes, the effects of inbreeding and genetic drift become more pronounced, which has negative impact on genetic diversity and thus contributes to the extinction vortex.

Genetic drift

Genetic drift is a random change in allele frequencies occurring in finite populations from generation to generation (Fisher 1930; Wright 1931). While
genetic drift occurs in all finite populations, it has major effects in small populations where it leads to faster fixation and loss of (rare) alleles. Loss of genetic variation via genetic drift has various fitness consequences.

In small populations, selection is less effective at purging slightly deleterious mutations because genetic drift leads to stochastic changes in allelic frequencies (Kimura 1957; Ohta 1992). This process can lead to an accumulation of deleterious alleles, increasing the genetic load and leading to a loss of fitness (Hedrick 2001). These changes can further decrease the population size, which in turn leads to further accumulation of deleterious mutations in a phenomenon known as mutational meltdown (Lynch & Gabriel 1990; Lynch et al. 1995b).

Loss of alleles via genetic drift also results in reduced standing genetic variation, which is the basis for adaptive potential and thus the ability of a population to respond to environmental changes and other challenges, for example competitors and diseases. Furthermore, fewer new mutations emerge in small populations, and (just as selection is less effective at purging slightly deleterious mutations) small populations are thought to be less effective at maintaining slightly advantageous mutations (Hedrick 2001). As a consequence, the potential for adapting to changing conditions is reduced (Kohn et al. 2006).

Inbreeding

Inbreeding, mating of related individuals, was associated with harmful effects already by Charles Darwin based on his experiments with plant mating systems (Darwin 1867; Charlesworth & Charlesworth 1987). Reduced fitness in offspring of related individuals is termed inbreeding depression, and this is a particularly severe problem in small populations where the opportunities for mating are limited and all individuals eventually become related after a number of generations (Keller & Waller 2002; Allendorf et al. 2013).

Inbreeding depression is a consequence of increased homozygosity, which affects fitness in two ways – partially recessive deleterious alleles are unmasked in homozygous state and loci with heterozygote advantage (overdominance) lose selective advantage (Charlesworth & Willis 2009).

Despite being of high concern for conservation biology, proper understanding of the genetic basis of inbreeding and the effects of inbreeding depression on population viability and growth have been hindered by the difficulty of studying inbreeding in wild populations (Kardos et al. 2016). Until a few years ago, detailed pedigrees requiring extensive observations were used
for estimating the inbreeding coefficient (Wright 1922). However, large-scale genomic datasets have enabled genome-based estimates of inbreeding, which facilitate studies of inbreeding in natural populations without pedigrees and provide a higher level of accuracy (Kards et al. 2016; Hedrick et al. 2017).

Conservation palaeogenomics

With an increasing amount of genomic data from non-model organisms of conservation interest, it is becoming clear that the relationship between genomic diversity and current population size is influenced by various factors, e.g. ancient bottlenecks and different life history strategies, and that genomic diversity therefore is not a consistent indicator of extinction risk (Diez-del-Molino et al., in press). To properly understand the genetic threats that a population has faced, the actual loss of variation can be quantified by including temporally distributed samples (utilizing museum or pre-historic specimens), which can define baseline levels of genetic variation before a bottleneck (Taylor et al. 1994). Furthermore, integrating temporal sampling and genome-wide diversity estimates on a population level can provide a fine-scale view of the microevolutionary processes operating throughout and after a bottleneck. In other words, the availability of large-scale data sets for non-model organisms and the possibility to track populations through time make it feasible to empirically test the role of extinction factors and their fitness consequences (Thomas 2012; Palkopoulou et al. 2015; Diez-del-Molino et al., in press).

Study system

The woolly mammoth (Mammuthus primigenius) is an excellent model system for studying pre-extinction microevolutionary processes. Originally an abundant species with widespread distribution, it faced a number of challenges at the end of Pleistocene, including numerous climatic shifts and hunting by expanding populations of modern humans (Martin & Stuart 1995; Stuart et al. 2004; Lister & Stuart 2008; Nikolskiy et al. 2011). By studying samples spread over the last 60,000 years of the woolly mammoth’s existence, we can track how a species, adapted to cold steppe-tundra habitat, dealt with increasing temperatures and associated environmental changes. Moreover, the last population that survived on Wrangel Island outlived the last mainland mammoths by at least 6000 years, making it a perfect model of a terminal
refugium for investigating the effects of genetic drift and inbreeding all the way to the species’ extinction.

The natural history of mammoths

Mammoths (genus *Mammuthus*), together with the present-day elephants, belong to the family Elephantidae, which originated in Africa in the Late Miocene and diverged into three lineages (*Loxodonta*, *Elephas*, and *Mammuthus*) that are clearly present in the fossil record from about 4-5 million years ago (Maglio 1973; Kalb & Mebrate 1993). Early genetic studies using short mitochondrial markers produced conflicting results regarding the phylogenetic relationships between mammoths and elephants (Hagelberg *et al.* 1994; Yang *et al.* 1996; Shoshani *et al.* 1998; Thomas *et al.* 2000). Later on, using complete mitochondrial genomes, the woolly mammoth was resolved as a sister lineage to the Asian Elephant (*Elephas maximus*; Krause *et al.* 2006; Rogalev *et al.* 2006; Rohland *et al.* 2007). According to molecular dating, the lineages of *Mammuthus* and *Elephas* diverged 6 million years ago (Brandt *et al.* 2012).

What is known about the mammoth’s biology is based on combining evidence from the remarkably preserved mammoth remains in permafrost, fossil assemblages, cave art depicting them, and insights from modern elephants (Lister & Bahn 2009).

Mammoths were large herbivores that, unlike most mammals but similarly to elephants, continued to grow throughout their lives. Because grasses and sedges dominated their diet, mammoths exchanged six sets of molars throughout life to replace the worn out enamel ridges (Lister & Bahn 2009). Based on analyses of isotopic patterns, mammoths are assumed to have formed family groups comparable to elephant herds (Hoppe 2004), consisting of a matriarch, adult females, and juveniles (Agenbroad & Mead 1987; Lister & Agenbroad 1994).

Evolutionary history of the woolly mammoth

The woolly mammoth has seen both its dawn and final demise in northeastern Siberia (Lister 1996b). It evolved from the steppe mammoth (*Mammuthus trogontherii*), with the morphological transition in the fossil record starting around 750,000 years before present (BP) and fully-evolved woolly mammoths present by at least 400,000 years BP (Lister & Sher 2001; Lister *et al.* 2005b; Lister & Bahn 2009). Thanks to the early development and stability of the periglacial steppe-tundra environment dominated by treeless grasslands (Sher *et
northeastern Siberia likely served as “an area of successive allopatric innovations” spreading to other parts of the northern hemisphere in times of favourable conditions (Lister & Sher 2001).

The woolly mammoth evolved morphological and physiological adaptations to the cold and open habitat that dominated the northern hemisphere during the Ice Age (Lister & Bahn 2009; Campbell et al. 2010; Lister 2014; Lynch et al. 2015). However, the onset of dramatic climatic fluctuations, starting with the Last Glacial Maximum (LGM) about 25,000 years BP, also marked the onset of a series of range contractions and expansions, which are thought to have resulted in an overall reduction of the mammoth population size (Lister & Stuart 2008).

These environmental shifts escalated in the Allerød interstadial period (14,000 to 12,900 years BP) when boreal birch and pine woodlands replaced the open steppe-tundra habitat and mammoths only survived in northern refugia, for example in the Taimyr Peninsula region (Stuart 2005; Lister & Stuart 2008).

Although mammoths re-expanded during the Younger Dryas (12,900 to 11,700 years BP), reaching as far as the East European Plain, shortly after the Pleistocene/Holocene transition they completely disappeared from the mainland (Stuart et al. 2002; Stuart et al. 2004). After approximately 10,000 years BP, there are only two populations known to have survived on two small islands. On Wrangel Island in the Chukchi Sea and Saint Paul Island in the Bering Sea, populations of woolly mammoths persisted until the mid-Holocene (Vartanyan et al. 1993; Vartanyan et al. 1995; Guthrie 2004). While there is good evidence that the mammoths on St. Paul Island disappeared when the main fresh water source on the island dried out approximately 5600 years BP (Graham et al. 2016), the mechanism behind the extinction of the Wrangel Island mammoths, which took place approximately 4000 years BP, is not so clear (Thomas 2012; Roca 2015).

**Wrangel Island mammoths**

With the end of the Ice Age, global sea levels rose due to increasing temperatures and melting ice sheets, which isolated a piece of elevated Siberian landscape, now known as Wrangel Island (Figure 1). The connection to the mainland was probably lost by 10,500 years BP when the land bridge to Chukotka was flooded (Vartanyan et al. 2008; Arppe et al. 2009).

In the Pleistocene, lush grasslands on what later became Wrangel Island are thought to have attracted mammoths during seasonal migrations and served as a summer feeding ground. However, when large parts of the coastal plains were
flooded in the Holocene, the island shrank to about 7600 km² in area. A large part of the island constitutes mountain ranges, inactive rock, perennial snowfields, and wetlands (Vartanyan et al. 2008; Arppe et al. 2009; Lister & Bahn 2009), and does consequently not represent an ideal refugium for a megaherbivore. The carrying capacity of Wrangel Island has been estimated to 149-819 mammoths based on Damuth’s equation using the relationship between body size and population density for herbivores in arctic environments (Nyström et al. 2010).

Based on a previous whole-genome analysis, the Holocene effective population size \( N_e \) was ~328 mammoths (Palkopoulou et al. 2015), which is far below the \( N_e \) of 500-5000 individuals considered in conservation biology to be required for long-term survival of populations (Franklin 1980; Franklin & Frankham 1998; Lynch & Lande 1998).

The first insights into the consequences of isolation on Wrangel Island were provided by analyses of ~750 base pairs (bp) of mitochondrial DNA (Nyström et al. 2010) and four nuclear microsatellites (Nyström et al. 2012). Both types
of genetic markers suggested a marked loss of genetic diversity due to an initial bottleneck around the time that the island was formed. Furthermore, genomic analyses comparing a Pleistocene mainland mammoth with one of the last Wrangel Island mammoths have showed a pronounced effect of inbreeding (Palkopoulou *et al.* 2015) and accumulation of detrimental mutations consistent with mutational meltdown (Rogers & Slatkin 2017) in the Holocene individual.

**Objectives**

The aim of this thesis was to use biomolecular methods to explore the biology and evolutionary history of the woolly mammoth. The extinction process was examined by studying genomic footprints of the events taking place during the last millennia of its existence.

More specifically, the aims were to:

- Reconstruct the demographic history of woolly mammoths as the population on Wrangel Island was established (*Paper I*).
- Test the theoretical assumption that balancing selection can maintain diversity throughout a bottleneck by comparing rates at which genetic diversity was lost at neutral loci and loci normally under balancing selection (*Paper II*).
- Investigate the molecular sex of woolly mammoth fossils to make predictions about its social structure (*Paper III*).
- Explore ways of obtaining complimentary socioecological patterns in extinct species by evaluating the potential of ancient hairs as a source of endogenous testosterone (*Paper IV*).
- Fill in gaps in the genomic timeline of the woolly mammoth extinction in order to provide insights into the consequences of genetic drift and inbreeding in small isolated populations (*Paper V*).
Methods

Type of material
The samples analysed in this thesis consist of bones, teeth, and tusks of the woolly mammoth collected in northeastern Siberia by our Russian collaborators. These samples have been freshly exposed from melting permafrost and are thus relatively well preserved.

All samples in this thesis were drilled using a multi-tool, hand-held Dremel drill in order to obtain fine bone/tooth powder. First, a thin layer was removed from the surface of the samples to avoid outer contamination, and then a hole was drilled in the sample until about 50mg of powder were retrieved.

Dating
All our samples have been radiocarbon dated, either as a part of previous research or through new direct dates that were obtained using Accelerator Mass Spectrometry (AMS) at the Oxford Radiocarbon Accelerator Unit and at the Centre for Isotope Research in Groningen. Throughout the study, the $^{14}$C dates are reported in conventional radiocarbon years before present (BP), including correction for isotopic fractionation and usage of conventional half-life (Mook & van der Plicht 1999). The dates were calibrated to calendar years using the IntCal13 calibration curve (Reimer et al. 2013) in the program OxCal 4.2 (Ramsey 2009) and medians of the calibrated/calendar years relative to 1950 CE (calBP) are used throughout the thesis.

Ancient DNA
Since the 1980s when the first studies using DNA of extinct species were published (Higuchi et al. 1984; Pääbo 1989; Thomas et al. 1989), the field of ancient DNA (aDNA) has been evolving rapidly with improving techniques and methods. Most notably, the introduction of the polymerase chain reaction (PCR) in the early 1990s, and the development of high-throughput sequencing
(HTS) techniques in the last decade, have revolutionized the analysis of ancient and historical samples.

Although the main interest in aDNA has concerned investigations of human ancestry and hominin relationships (e.g. Noonan et al. 2006; Briggs et al. 2009; Green et al. 2010) and phylogenetic relationships of extinct species (e.g. Thomas et al. 2000; Barnes et al. 2002; Shapiro et al. 2004; Lister et al. 2005a), the power to look back in time has found many other applications, for example in studies of demography, admixture, selection, and genetic basis of phenotypic variation (Millar & Lambert 2013; Shapiro & Hofreiter 2014). Not only has palaeogenetics transitioned to palaegenomics, we are now able to do population-level studies with complete genomes that are many thousands of years old.

Despite the current achievements, the early advances of aDNA research were hindered by a number of issues specific for the work with ancient DNA molecules, especially dealing with contamination and DNA damage (Pääbo et al. 2004).

**Contamination precautions**

Ancient samples are exposed to various sources of contamination by exogenous DNA, something that became especially obvious when scientists started using shotgun sequencing, i.e. non-targeted sequencing of the total DNA content. Shotgun sequencing of Neanderthal samples (Green et al. 2010) showed that 95-99% of the DNA belonged to a non-primate source. These contaminants mostly consist of microbes present in the environment where the organism was buried after death and bacteria participating on decomposition of the dead tissue. A genomic analysis of exogenous DNA from a cave bear sample showed that 66% of the DNA did not match any record in nucleotide databases (Noonan et al. 2006), demonstrating the gaps in microbial genetic sampling.

Another source of contamination is human DNA from people excavating, collecting and handling the samples, as well as cross-contamination from other samples with higher proportion of endogenous (i.e. authentic) DNA. Several studies from the early days of aDNA research have later been believed or even proved a result of contamination, receiving the label “antediluvian DNA” (Lindahl 1993b). This led researchers to generate lists of criteria for aDNA studies to be considered authentic (Pääbo et al. 1989; Cooper & Poinar 2000). However, these were later on criticized because rather than rationally evaluating risks on a case-to-case basis, scientists relied on a generalized checklist (Gilbert et al. 2005a).
To avoid contamination, the laboratory work described in this study was performed in a dedicated clean aDNA laboratory at the Swedish Museum of Natural History, where modern DNA has never been handled. The lab is equipped with a HEPA-filter system and is physically separated from the post-PCR laboratories. Inside the aDNA lab, standard procedures such as using whole-body suits, face masks, and double layers of gloves are followed. The surfaces are regularly sterilized with sodium hypochlorite and the hoods designated for work are UV-lighted. All tools are sterilized in a UV crosslinker before and after use. Negative controls are used at all steps of laboratory work to detect contamination. Furthermore, using mapping quality filters and other approaches, contamination can usually be recognized and removed bioinformatically.

DNA damage

Post-mortem damage is a specific feature of historical and ancient DNA. After the death of an organism, DNA is affected by enzymatic damages caused by endonucleases, which are no longer removed by DNA repair mechanisms and lead to changes in the nucleotide sequence (Lindahl 1993a; Pääbo et al. 2004). Among the three most common types of DNA damage - hydrolysis, oxidation, and Maillard reactions – hydrolytic decomposition is the main problem. Hydrolysis affects sensitive glycosidic base-sugar bonds in DNA in two ways: either by direct cleavage and depurination leaving abasic sites susceptible to strand breaks, which results in fragmentation into short pieces, or by introducing miscoding lesions, most often observed as deamination of cytosine to uracil at the 5’-ends of molecules (Lindahl 1993a; Gilbert et al. 2005b; Binladen et al. 2006).

In this thesis, the vast majority of samples were treated with uracil-DNA-glycosylase (UDG), which removes uracils and leaves abasic sites that are cleaved by endonuclease VIII, introducing strand breaks in place of nucleotide misincorporations (Hofreiter et al. 2001; Briggs et al. 2010). This considerably reduces the amount of DNA damage in genomic data.

Moreover, DNA damage was considered during bioinformatic analyses by means of applying more conservative approaches during consensus calling, for example, of mitochondrial genomes, where only positions covered by at least three individual reads were called, or in the pairwise sequentially Markovian coalescent (PSMC) analysis of complete genomes, where we excluded sites with a depth below 1/3 of the average coverage of each genome, which was at
least 8X (i.e. not allowing positions with less than two or three reads to be called).

DNA extractions

Most DNA extraction methods used in ancient DNA today are based on binding DNA to silica (Höss & Pääbo 1993; Yang et al. 1998; Hofreiter et al. 2004; Rohland & Hofreiter 2007; Dabney et al. 2013). Almost all samples in this thesis have been extracted following the Yang et al. (1998) protocol, with modifications as described in Ersmark et al. (2015).

The bone powder is mixed with the extraction solution, consisting of ethylene diaminetetra-acetic acid (EDTA), proteinase K, and Urea, and is then incubated overnight under motion. The supernatant is subsequently concentrated using a Vivaspin® (Sartorius) filter and mixed with PB buffer, which helps DNA bind to the silica column. In the next step, the DNA bound to the silica column is washed using PE buffer and eluted in 100 μl of EB buffer (QIAquick PCR Purification Kit, QIAGEN).

For nine samples, a second set of extractions was performed using the Dabney et al. (2013) protocol specifically adapted for degraded samples with short fragment lengths. This protocol, which has been used to extract DNA from a Middle Pleistocene cave bear, is based on the standard silica protocol, but instead of PB buffer uses a custom binding buffer (guanidine hydrochloride, isopropanol, sodium acetate, and Tween-20) and uses a larger volume of the binding buffer. However, the endogenous DNA yield using this protocol was not markedly improved compared to the standard protocol described above (unpublished data). Throughout the thesis, the only genetic data based on the Dabney et al. (2013) protocol is from E469D, the last woolly mammoth known to science.

PCR

Since its early days (Pääbo & Wilson 1988; Hagelberg et al. 1989; Pääbo 1989; Hagelberg et al. 1994), the aDNA field depended on amplification by PCR because endogenous molecules in ancient samples are in small quantities. Compared to standard PCR protocols, there are several modifications adjusted for the specifics of aDNA. First, a high-performing DNA polymerase able to handle DNA damage is preferred (Fulton & Stiller 2012). Second, bovine serum
albumin (BSA) is often added to the PCR mastermix in order to bind PCR inhibitors, which are abundant in aDNA samples and interfere with DNA polymerase (Eilert & Foran 2009; Woide et al. 2010). Third, a higher number of amplification cycles are generally needed in aDNA studies due to the low concentrations of ancient molecules.

In **Paper II**, PCR with tagged primers (Binladen et al. 2007) was used to generate barcoded PCR products that were pooled and sequenced in a single run on a GS Junior (454 Life Sciences). Considering the previously described issue with DNA damage, two independent PCR replicates were generated for each sample. Moreover, for all homozygotes a third replicate sequenced by Sanger sequencing was added to minimize the risk that homozygous genotypes were the result of allelic dropout.

**High-Throuhgout Sequencing**

Two types of high-throughput sequencing (HTS) were used in this thesis, Roche 454 pyrosequencing (Margulies et al. 2005) and Illumina sequencing (Bentley et al. 2008), both representing the second generation of HTS and both being based on the “sequencing-by-synthesis” principle.

In the early phase of HTS, the main strengths of the Roche 454 method were mainly long read length and short run time, whereas Illumina’s advantages were lower cost and larger data output. While the 454 platform has been discontinued since 2016, Illumina continues to increase its range of sequencers.

**Roche 454 pyrosequencing**

In **Paper II**, amplicon sequencing of 4 loci of the major histocompatibility complex (MHC) was performed using the Roche 454 pyrosequencing method. Newly designed primers with unique barcodes on 5’ and 3’ ends (Binladen et al. 2007) were used to analyse 32 samples, with 2 replicates each, simultaneously. A total of 181 amplified PCR products were pooled and sequenced on an in-house GS Junior platform (454 Life Sciences), following the protocols supplied with the sequencer.

In **Paper IV**, the sequence data (Gilbert et al. 2007; Gilbert et al. 2008) had been generated by whole-genome shotgun sequencing using the Roche 454 pyrosequencing method, on a Roche GS FLX sequencer.
Illumina shotgun sequencing

Data in Papers I, III, and IV was produced using whole-genome shotgun sequencing on the Illumina platform at the Science for Life Laboratory in Stockholm (project number b2015028). Sequencing in Papers I and III was performed on the Illumina HiSeq2500, while sequencing in Paper V was done on both the HiSeq2500 and HiSeqX.

Illumina libraries were built following the Meyer & Kircher (2010) protocol for paired-end, multiplexed sequencing, including uracil-DNA-glycosylase treatment to remove uracils.

Bioinformatics

Data processing

In Paper II, the following set of tools was used to demultiplex samples, trim the adapters, and sort the 4 loci: 454 Sequencing System Software 2.9 provided with the GS Junior, a perl script fastaQual2fastq.pl (available from https://github.com/josephhughes/Sequencemanipulation/blob/master/fastaQu al2fastq.pl), and Geneious 7.0.3 (Kearse et al. 2012). In Papers I, III, and V, SeqPrep 1.1 (available from https://github.com/jstjohn/SeqPrep) was used to trim Illumina adapters and to merge paired-end reads. In Paper IV, cutadapt (Martin 2011) was used to trim the adapters.

Throughout the thesis, sequencing data was processed using a modified version of the bioinformatic pipeline by Clarke et al. (2014), utilizing BWA (Li & Durbin 2009, 2010) for mapping and SAMtools (Li et al. 2009) for processing the alignment. In Paper II, the data was mapped to the Asian elephant MHC reference sequence (GU369701; Archie et al. 2010), while in Papers I, III, IV, and V, the reads were mapped against a merged reference consisting of the nuclear genome of an African savannah elephant (LoxAfr4; Broad Institute) and mitochondrial genome of a woolly mammoth (Krause; DQ188829; Krause et al. 2006).

Data analyses

In Paper I, a haplotype network of 21 newly generated and 21 previously published mitogenomes was constructed in PopART (available from http://popart.otago.ac.nz) and a dated phylogeny was reconstructed in BEAST 1.8.0 (Drummond et al. 2012). The evolution of female effective population sizes was plotted in a Bayesian Skyride analysis and mutation rates were
estimated both in BEAST and using coalescent simulations in fastsimcoal 2.5.2.2 (Excoffier et al. 2013) and arlsumstats 3.5.2 (Excoffier & Lischer 2010), controlled by custom R scripts (R Development Core Team 2013).

In Paper II, changes in genetic diversity in response to the bottleneck were evaluated by comparing 12 mammoths from before and 12 mammoths from after the isolation on Wrangel Island. Genetic diversity was quantified as expected heterozygosity, observed heterozygosity, number of alleles, and number of unique alleles in Arlequin 3.5.2.1 (Excoffier & Lischer 2010). Ratios of non-synonymous and synonymous mutations were used to scan for selection using MEGA6 (Tamura et al. 2013). To further test for selection, coalescent simulations were performed using the softwares described above.

In Papers III and IV, the biological sex of mammoth samples was determined by estimating the ratio between the number of reads mapping to chromosomes X and 8 in the reference genome, while normalizing for the size of the chromosomes. The expected ratio is ~0.5 in males and ~1 in females. Confidence intervals were estimated from the standard deviation of unambiguously determined samples.

In Paper V, temporal changes in genome-wide heterozygosity were estimated using six complete mammoth genomes in two ways, using genotype hard calls generated with SAMtools (Li et al. 2009) and bcftools (Li 2011), and using a likelihood-based method in mlRho (Haubold et al. 2010). Runs of homozygosity (ROH) were identified as homozygous fragments above 0.5MB in plink 1.9 (available form https://www.coggenomics.org/plink2). Changes in effective population size through time were reconstructed using the PSMC method (Li & Durbin 2011).
Results & Discussion

Demographic history of the Wrangel Island mammoths

Mitochondrial DNA has been a popular genetic marker for investigating the woolly mammoth’s evolutionary history (Barnes et al. 2007; Gilbert et al. 2007; Debruyne et al. 2008; Gilbert et al. 2008; Palkopoulou et al. 2013). The Late Pleistocene mammoth population seems to have been diverse and stable in female effective population size (Barnes et al. 2007; Nyström et al. 2010; Palkopoulou et al. 2013), but most of the mitochondrial diversity was lost when mammoths became isolated on Wrangel Island (Nyström et al. 2010).

In Paper I, using 42 Pleistocene and Holocene mammoth mitogenomes, I investigated the extent and consequences of changes in genetic diversity during the isolation of Wrangel Island. In the Pleistocene mammoths, 27 different haplotypes were identified, whereas only seven haplotypes were observed in Holocene samples from Wrangel Island (Figure 2). The seven Wrangel haplotypes form a separate subgroup, shaped in a star-like pattern, which suggests that the population was established by a single ancestral haplotype. This result is supported by the Bayesian phylogeny, where Wrangel Island mammoths form a monophyletic group. Central Siberian mammoths are basal to the Wrangel Island cluster in the phylogeny, which suggests that they might be ancestral to the Wrangel Island mammoths. Central Siberia, specifically the Taimyr Peninsula, has been recognized as a refugium based on the continuity of radiocarbon dates, which suggests a continuous population through the end-Pleistocene climatic oscillations (Stuart et al. 2002; Lister & Stuart 2008). Considering that Wrangel Island mammoths are more closely related to Taimyr samples even compared to end-Pleistocene Wrangel and Pevek region samples, a plausible explanation is that along with the westward re-expansion from Taimyr during Younger Dryas (Stuart et al. 2002; Lister & Stuart 2008), mammoths re-expanded also eastwards, finding a Holocene refugium on Wrangel Island.
Figure 2. Holocene Wrangel Island mammoths cluster together in the Bayesian phylogeny (A) and form a subgroup shaped in a star-like pattern in the median-joining haplotype network (B). More details in Paper 1.

The reduction of mitogenome diversity down to a single lineage during the time that Wrangel Island was formed suggests that a founder effect took place when the Wrangel population was established. In line with that, Bayesian Skyride analysis showed a 15-fold decrease in female effective population size during the Pleistocene/Holocene transition, consistent with a previous estimate (Figure 3; Palkopoulou et al. 2013).

Figure 3. Bayesian Skyride plot of the female effective population size ($N_e$) scaled by a factor of generation time (purple line, left y-axis), laid over the climate record (black line, right y-axis).
Based on the Bayesian Skyride plot (Paper I) and PSMC analyses (Paper V; Palkopoulou et al. 2015), the effective population size on Wrangel Island was small (~300-500 individuals), but remained stable until the extinction 4000 years BP.

Functional consequences of the bottleneck

Detrimental variation

In Paper I, I found that among three synapomorphies in the mitogenomes of Wrangel Island mammoths, one is a nonsynonymous mutation in a highly conserved region of the ATP6 gene, which is a gene encoding for a subunit of the ATP synthase enzyme of the oxidative phosphorylation pathway. While this mutation did not occur in any of the Pleistocene mammoths, it was present in all Holocene Wrangel individuals. Taking into account that the mutation occurred in a conserved region of a functionally important gene and mutations at neighbouring sites give rise to disease phenotypes, it is likely that this mutation was deleterious. Random fixation of a detrimental allele due to genetic drift at a small population size (Ohta 1992; Lynch et al. 1995a) is in line with the reduction in effective population size associated with the founder effect.

Adaptive variation

To be able to respond to new environmental challenges, populations need to maintain adaptive genetic variation. MHC is a good example, because it is a gene family responsible for pathogen resistance (Piertney & Oliver 2006). In order to recognize a wide spectrum of pathogens, balancing selection maintains a high diversity of MHC alleles (Penn et al. 2002; Spurgin & Richardson 2010; Oliver & Piertney 2012).

In Paper II, the mammoth MHC diversity was analysed before and after the isolation on Wrangel Island in order to evaluate if balancing selection maintained a stable level of genetic diversity throughout the bottleneck. I compared 12 mammoths from each of the Pleistocene and Holocene populations. The average observed heterozygosity in the three analysed fragments of the MHC class II DQA locus decreased by 42% (Table 1) following the bottleneck. The average number of alleles dropped by 37%, and while there were on average 2.3 unique alleles in the Pleistocene population, no unique alleles were present in Holocene (Table 1). This loss of
adaptive variation was comparable to the decrease in neutral variation at four microsatellite loci observed in a previous study (Nyström et al. 2012), where the average observed heterozygosity decreased by 31%.

Table 1. The estimates of genetic diversity before and after the isolation on Wrangel Island. $H_o =$ observed heterozygosity; $H_e =$ expected heterozygosity; $N_a =$ number of alleles.

<table>
<thead>
<tr>
<th>Locus</th>
<th>&gt; 13 000 yr BP (n=12)</th>
<th>&lt; 10 000 yr BP (n=12)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$H_o$</td>
<td>$H_e$</td>
</tr>
<tr>
<td>Exon 2</td>
<td>0.83</td>
<td>0.85</td>
</tr>
<tr>
<td>Intron 2</td>
<td>0.92</td>
<td>0.75</td>
</tr>
<tr>
<td>Exon 4</td>
<td>1.00</td>
<td>0.80</td>
</tr>
<tr>
<td>Mean</td>
<td>0.92</td>
<td>0.80</td>
</tr>
<tr>
<td>± SD</td>
<td>±0.08</td>
<td>±0.05</td>
</tr>
</tbody>
</table>

Although the scans of selection did not find significant signals (possibly due to the small sample size), the trans-species polymorphism (Figure 4), sharing of MHC alleles between species (in this case with elephants; Archie et al. 2010), is a signal of long-term balancing selection (Takahata & Nei 1990).

![Figure 4. Alignment of predicted amino acid sequences of the MHC DQA exon 2, showing the trans-species polymorphism in the woolly mammoth (Mam), the African savannah elephant (Lod), and the Asian elephant (Elm). The highlighted positions indicate predicted antigen-binding regions in the elephant MHC (Archie et al. 2010).](image)

As an alternative way to test for the roles of selection and drift during a population bottleneck, coalescent simulations were used to infer whether the
observed loss of genetic diversity on MHC can be explained solely by genetic drift. The results suggest that under a wide range of effective population sizes at the time of the bottleneck and using previous estimates of Holocene population sizes (Nyström et al. 2010; Nyström et al. 2012; Palkopoulou et al. 2015), the loss of 37% of alleles can be explained solely by the reduction in population size. Hence, there was no evidence that selection maintained MHC diversity during and after the bottleneck that took place as mammoths became isolated on Wrangel Island.

Behaviour and social structure

In Paper III, biological sex of 98 woolly mammoths was analysed using low-coverage genomic data and by comparing the proportions of reads mapping to chromosomes X and 8. Sex was successfully assigned to 95 of the analysed mammoths and a biased sex ratio was observed, with 69% of the samples representing males (Figure 5). Considering that a similarly skewed sex ratio has been observed among mammoth (including the Columbian mammoth, *Mammuthus columbi*) and mastodon (*Mammut*) remains in fossil assemblages from natural traps (Coope & Lister 1987; Lister & Agenbroad 1994; Lister 2009), we suggested a hypothesis that the excess of males is a consequence of male individuals being more likely to die in such natural traps, which ensure their preservation.

The hypothesis proposed by Agenbroad & Mead (1987; 1994) to explain the male-biased sex ratio in the Hot Springs sinkhole seems like a plausible explanation. This hypothesis is based on the behaviour and social structure of mammoth family groups. Just as elephants, mammoth herds were likely dominated by a matriarch and her offspring (Hoppe 2004;
Young adult males probably had to leave the group, and lived either roaming the steppe-tundra alone or with other bulls in bachelor groups. Solitary males would thus have been more likely to enter dangerous grounds or end up in a natural trap than careful females with calves (Agenbroad & Mead 1987; Lister & Agenbroad 1994; Lister 1996a). As described by Lister & Agenbroad (1994): “One can imagine a Gary Larson cartoon depicting a group of young bachelor mammoths gathered at the edge of the Hot Springs sinkhole 26,000 years ago, daring each other to traverse the steep walls for forage or water. Any such adventurous trip – for whatever reason it was undertaken – was a one-way trip. The age analysis of the mammoth population at Hot Springs fits such a model”.

Another approach of obtaining more information about life history strategies of extinct species is described in Paper IV. Presence and quantity of testosterone was measured in hair shafts of ten woolly mammoths (Gilbert et al. 2007; Gilbert et al. 2008) using liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS; Koren et al. 2002; Koren et al. 2012). Testosterone was successfully measured in all ten samples, and a comparatively high variation was observed among individuals with testosterone levels ranging from 0.77 to 4.57 pg/mg hair. When the genome-based sexing analysis was applied to the samples, a higher mean level of testosterone was observed in males compared to females. Although the sample size was too small for statistical testing, these results open up for novel opportunities to investigate biological properties in ancient samples. Since testosterone is involved in social processes, for example social status formation, mating, and risk-taking behaviour, measuring testosterone levels in a larger number of individuals could provide insights into socioecology of species.

Genome erosion

In Paper V, four complete genomes were generated at medium-high coverage, which together with two previously published genomes (Palkopoulou et al. 2015) span the last 45,000 years of the woolly mammoth’s evolution. This time period encompassed the range shifts at the end of Pleistocene (during the Bolling-Allerød and the Younger Dryas), the Pleistocene/Holocene transition, and the isolation on Wrangel Island.

Based on average autosomal heterozygosity, the main loss of genome-wide diversity occurred at the Pleistocene/Holocene transition, when mammoths lost ~30% of their genome-wide diversity (Figure 6). During the Holocene,
decline continued at a slower rate as suggested by a 6%-decrease between ~7300 and ~4300 years BP. However, the last mammoth, only ~300 years younger than the previous one, had lost an additional 20% of genome-wide diversity.

![Figure 6. Genome-wide heterozygosity in six complete mammoth genomes plotted against time and climatic record from the North Greenland ice core.](image6.png)

The proportion of the genome in runs of homozygosity (ROH) was 4-fold higher in the Holocene than in the Pleistocene, suggesting inbreeding in the Wrangel Island population. During approximately ten generations, the last mammoth acquired an additional 12% of ROH (Figure 7). As a result, 38.5% of the last known woolly mammoth’s genome was completely depleted of

![Figure 7. Cumulative size of runs of homozygosity showing the accelerated genome erosion during the ten generations separating the last two woolly mammoths.](image7.png)
diversity. Possibly, this might have been a result of intensified inbreeding as the effective population size became critically low.

Despite the steep decline in genomic diversity after the initial bottleneck and continued decline in Holocene, the genome erosion thus accelerated remarkably during the final three centuries. This is consistent with expectations from the extinction vortex model predicting that populations decline faster closer to extinction (Gilpin & Soulé 1986; Fagan & Holmes 2006).
Concluding remarks

Although the final cause of the woolly mammoth’s extinction remains yet to be clarified, insights from previous studies filled in important pieces of the puzzle of the woolly mammoth’s extinction (Thomas 2012; Roca 2015). Depending on the questions asked and genetic markers used to answer them, they offered different results. While the mtDNA and microsatellite analyses (Nyström et al. 2010; Nyström et al. 2012) highlighted the impact of the initial bottleneck but found stable levels of genetic variation until extinction, the whole-genome analyses found signs of genome erosion in one of the very late Wrangel mammoths (Palkopoulou et al. 2015; Rogers & Slatkin 2017). In this thesis, I explored these topics further and my results indicate that rather than contradictory, previous results provide complementary information, revealing the complexity of processes that took place on Wrangel Island after its separation from Siberia.

The results from mitogenomes and MHC provide support for the severe bottleneck associated with the establishment of the Wrangel Island population (Papers I and II). Moreover, the loss of genetic diversity at the Pleistocene/Holocene boundary was here also quantified using genome-wide heterozygosity estimates (Paper V) and the decrease seems even greater than indicated previously (Palkopoulou et al. 2015). However, reduced heterozygosity in bottlenecked populations does not necessarily lead to extinction, as suggested by examples of populations surviving with extreme levels of genomic homozygosity (Dobrynin et al. 2015; Robinson et al. 2016). Indeed, for three thousand years (separating the two earlier Wrangel mammoths) the decline in genome-wide heterozygosity was rather moderate and could be associated with the seeming stability of the population as reported at mtDNA and microsatellites (Nyström et al. 2010; Nyström et al. 2012). Nevertheless, the extraordinary tempo of decline in the last three hundred years (Paper V) is theoretically consistent with mutational meltdown (Lynch et al. 1995b), with predictions from the extinction vortex model (Caughley 1994) and empirical observations from final declines of ten wild vertebrate populations (Fagan & Holmes 2006). Beyond these results, one can only speculate about the
causes and consequences of the fast genomic deterioration just before the extinction. There are two main hypothesis to be considered: 1) could background levels of inbreeding and accumulation of deleterious variation due to a long-term small population size be responsible for the population reaching some sort of a tipping point, or 2) was the intensified genomic deterioration a result of a sudden change, for example in the environment or in the presence of humans, which disrupted a fragile equilibrium.

The link between genetic diversity, population size, and extinction risk is complex and is influenced by a number of factors related to the species’ biology. For example, mito-nuclear incongruence has been described in all contemporary species of elephants, and in the Columbian and woolly mammoths; and it has been attributed to the male-biased dispersal (Roca 2015). The social structure and behaviour in extinct species can be explored using molecular methods (Papers III and IV), revealing patterns that provide complementary insights into the inference of population dynamics of threatened and extinct species.

Until a sufficient number of conservation palaeogenomic studies accumulates and starts revealing more general patterns, cross-species comparisons of genetic diversity should be considered with caution (Diez-del-Molino et al., in press). Nonetheless, the last population of the woolly mammoth is an extraordinary model system, which can be used for testing some of the principal theoretical questions in population and conservation genetics. A detailed sampling of the Holocene Wrangel Island population might reveal a fine-scale picture of the pre-extinction processes that this thesis provided some clues about, for example regarding the accelerated evolutionary rate (Paper I), accumulation of detrimental variation (Paper I and V), loss of adaptive variation (Paper II), and genomic erosion (Paper V).

Using a fine-grain temporal genomic sampling, it might be possible to pinpoint these processes to particular events. Sampling through the bottleneck would provide insights into the origins of the Holocene Wrangel mammoths and into the extent of diversity in the population immediately after the isolation. Despite the attempt to sample the early phase of Wrangel Island isolation, all samples from this time period had consistently low endogenous DNA levels – a pattern that itself deserves more investigation. Finally, a thorough analysis of the last few hundred years on Wrangel Island will be essential for identifying what triggered the genome erosion and for connecting the changes in genetic diversity to fitness consequences, which could allow a better estimation of the extinction risk.
References


Gilbert MTP, Shapiro B, Drummond A, Cooper A (2005b) Post-mortem DNA damage hotspots in Bison (Bison bison) provide evidence for both damage and mutational hotspots in human mitochondrial DNA. Journal of Archaeological Science 32, 1053-1060.


Nikolskiy PA, Sulerzhitsky LD, Pitulko VV (2011) Last straw versus Blitzkrieg overkill: Climate-driven changes in the Arctic Siberian mammoth
population and the Late Pleistocene extinction problem. *Quaternary Science Reviews* 30, 2309-2328.


Stuart AJ (2005) The extinction of woolly mammoth (Mammuthus primigenius) and straight-tusked elephant (Palaeoloxodon antiquus) in Europe. Quaternary International 126, 171-177.


Svensk sammanfattning

utdöda arter är att använda molekylära metoder för att jämföra hormon-nivåer hos honor och hanar, vilket utnyttjades här genom att mäta mängden testosteron i mammuthårstrån. Slutligen sekvencerades hela den nukleära arvsmassan från fyra mammutar för att undersöka de genetiska konsekvenserna i den lilla populationsstorleken på Wrangels ö. Då dessa arvsmassor jämfördes med arvsmassor från den talrika istida fastlandspopulationen fann jag att mammutarerna på Wrangels ö hade lägre nivåer av genetisk diversitet och att en större del av deras arvsmassor bestod av långa segment helt utan genetisk variation (så kallade runs of homozygosity), vilket indikerar inavel. Värt att notera är att det utifrån dessa fynd framstår som att förlusten av genetisk variation accelererade under de tio sista generationerna före utdöendet, vilket resulterade i att nästan 40 % av arvsmassan hos den sista kända mammuten var helt utan diversitet. Resultaten från denna avhandling visar hur genetisk drift och inavel ledde till en utarmning av arvsmassan hos den sista överlevande populationen av ullhåriga mammutar. Trots att Wrangels ö var ett refugium där mammutar överlevde under flera tusen år efter istidens slut, och trots att de faktiska orsakerna för det slutgiltiga utdöendet ännu inte är klarlagda, så tyder dessa resultat på att isoleringen och den begränsade populationsstorleken utan möjligheter till inflöde av nytt genetiskt material kan ha lett till en reducerad livskraftighet och därmed också bidragit till mammutens utdöende.
Zhrnutie v slovenčine

Druhy na celom svete podliehajú zmenšujúcej sa početnosti a veľkosti areálu a ich osud v meniacom prostredí bude možno závisieť od schopnosti prežiť v malých a fragmentovaných populáciiach. Potreba pochopiť ako vymieranie funguje naberá na naliehavosti, ale naše poznatky o procesoch predchádzajúcich vyhynutiu sú stále obmedzené, pretože metódy umožňujúce skúmať populácnú a ochranársku genomiku u ohrozených druhov vo voľnej prírode, sa stali dostupnými iba nedávno. V tejto dizertačnej práci som používala poslednú populáciu mamuta srstnatého (*Mammuthus primigenius*) ako model na štúdium populánej dynamiky, ktorá predchádza vyhynutiu. Starobylú DNA som používala ako prostriedok na skúmanie mikroevolučných procesov v reálnom čase. Analyzovala som genetické zmeny, spôsobené environmentálnymi zmenami na konci poslednej doby ľadovej a zaobrala som sa predovšetkým dôsledkami genetického drifu a inbreedingu (príbuzenského kríženia), súvisiacimi s izoláciou mamutov na Wrangelovom ostrove, kde pri nízkej početnosti prežili až 6-tisíc rokov. Pomocou mitochondriálnych genómov som našla známky efektu zakladateľa, ktorý v čase, keď mamuty uviazli na Wrangelovom ostrove (pred zhruba 10 500 rokmi), znížil materskú diverzitu na jedinú líniu. Navyše, dvoj- až trojnásobný nárast mutačnej rýchlosti a prítomnosť’ fixedovej škodlivé mutácie v géne *ATP6*, ktorý kóduje jeden z klúčových enzymov oxidatívnej fosforilácie, je v súlade s predpokladom, že selekcia je menej efektívna v odstraňovaní škodlivých mutácií v malých populáciách. U imunitného génu, ktorý patrí medzi bielkoviny hlavneho histokompatibilného komplexu (MHC glykoproteiny), bola tiež pozorována strata diverzity, a to aj napriek tomu, že MHC je považované za gén pod vplyvom balansujúcej selekcie. Genomické dáta s malým pokrytím boli použité na analýzu podielu endogénej (autentickéj) DNA a pohlavia mamutích vzoriek. Pozorovanie vychýleného pomerenia pohlavia v prospech samcov (69%) nás viedlo k záveru, že mamutí samci boli náchylnejší zomrieť spôsobom, ktorý im zabezpečil lepšie uchovanie. Iný potenciálny spôsob získania informácií o socioekológii vyhynutých organizmov, ktorý sme tu skúmali, je prostredníctvom merania hladiny testosterónu v srsti mamutov v spojení
s určením pohlavia. Na záver, predchádzajúce výsledky naznačujú, že v Holocéne bola efektívna veľkosť populácie mamutov na Wrangelovom ostrove veľmi obmedzená a populácia bola možno príliš malá na to, aby sa vyhla genómovej erózii. Aby sme preskúmali dôsledky malej veľkosti populácie, osekvenovali sme štyri mamuty do relatívne vysokého pokrytia. Keď sme mamuty z holocénnego Wrangelovho ostrova porovnali s pôvodnou populáciou z Pleistocénu, zistili sme, že mamuty na Wrangelovom ostrove mali nižšiu celogenómovú diverzitu a vyšší podiel ich genómov pozostával z tzv. runs of homozygosity, čiže dlhých úsekov úplne zbavených variability. Čo je obzvlášť dôležité, genómová erózia zdanlivo zrychlila v priebehu posledných desiatich generácií, čoho následkom bolo, že genóm posledného mamuta známeho vede pozostával zo 40% z úsekov bez diverzity. Celkovo, naše výsledky zdôrazňujú, ako genetický drift a inbreeding spustili genetické chátranie v poslednej populácii mamuta srstnatého. Napriek tomu, že Wrangelov ostrov bol pre mamuty refúgiom, v ktorom prežili tisíce rokov po konci doby ľadovej, a príčinné faktory ich konečného vyhynutia ešte celkom nepoznáme, izolácia a malá veľkosť populácie bez akéhokoľvek toku génov možno prispeli k vymretiu.
Acknowledgements

First and foremost, I want to thank my supervisor Love Dalén for being the best kind of supervisor one can possibly imagine. There are so many things that I am thankful for that I don’t even know where to start. Maybe at the beginning then. Love, thank you for giving me a chance to do the most awesome PhD project ever. Thanks for always having a gazzilion of ideas, for having plans B, and for never once in my PhD letting me feel like this whole thing might not be successful. I have learnt so much and you have always been there to provide advice, to give me a high-five for all my small achievements, and to lead me through the challenging moments, which working in science necessarily brings along. Thank you for giving me so many different opportunities, like bringing me to Wrangel Island – even though it initially hadn’t been part of the plan, you still made it happen; and if I ever take over Igor’s job of driving Argo across Wrangel Island, I promise to send you all the samples I find. Which leads me to the most important part, thank you for making these four-and-something years such a fun experience, for all the jokes, for all the times you made fun of me (and I’m sure the video of me sawing my first mammoth tusk will always be a good way to blackmail me), also for understanding my sense of humour, for never getting angry (although we’ve got pretty close when you started using excessive amounts of exclamation marks on Slack), for your skill in surrounding yourself with great people, and you know, for being my Jedi Master.

Speaking of great people, I want to thank everybody who has belonged to Love’s group throughout the course of my studies – Vendela, Elle, Jean-Luc, Erik, Edson, Matti, Yvonne, George, Victor, Jonas, David, Johanna, Karin, Nic, and Tatiana – it is a privilege to call you my friends and I very much value the time we spent together. Elle, thank you for being my role model, for paving the way for the methods and pipelines that I’ve been using all the way through the PhD, for answering my never-ending questions, and last but not least, for introducing me to Les Mills. Vendela, Jean-Luc, and Erik, thank you for being my older PhD-siblings, you’ve been a great source of support, advice, and help. Matti, your passion for science is inspiring and I’m looking forward to all the
future email essays. Tatiana and Nic, I hope there will be more weekends with fika and sauna that end up getting published in Current Biology. Johanna, thank you for letting me play the role of senior PhD student, but more importantly, thank you for being such a kind and empathetic person. David, you deserve a special thank you, because you walked through this Mordor with me, and we all know that one does not simply walk into Mordor (this is a Dani-level joke as a gift for you). You and your scripting magic took the mammoth project several levels higher and that this PhD thesis is happening, is a great deal thanks to you. Even though you’re a Renegade, it was an honor to have you on my side.

On that note, I would like to thank my gaming squad – the Lovers, Dani, Eric, Miroslav, Allison, Victor (aka the Fredrikers), and Mozes. If not for all those hours wasted playing board games, finishing this thesis would have been much less stressful. And yet it was absolutely worth it, because nothing maintains mental health better than a few rounds of games where your colleagues become mortal enemies; and because of the loads of fun we had. I’m going to miss the alien kitten, the Granny, and Dani Lafayette, but most of all, I’m going to miss you.

I am also thankful to everybody at the museum, for making it a workplace worth getting out of bed for every morning. I want to thank Fredrik, Pia, Jane, and Anna-Lena, for being so kind and helpful with everything I ever needed. Anna-Lena, thank you for the lovely trip to Bromma to buy a drill, and all the other times you fixed stuff for me, always with a smile. I am also thankful to the people in the lab – Martin, Bodil, Rodrigo, Veronica, Niclas, Rasa, and Wendy – for helping me find stuff, show how things work, discuss why things don’t work, and keep the ancient DNA lab stocked up. Thank you to Johan for being the person to go to with bioinformatic trouble.

In parallel, I want to thank everybody at Zootis. Despite my sporadic appearances, you always made me feel welcome and a part of the team. Thanks to Anette, Siw, Minna, Birgitta, and Linda everything always went smoothly. Gabriella and Chris, my follow-up committee, thank you for helping to keep me on the right track. Sven, thanks for the invitations to Blodbadet – I very much enjoyed the company, the food, the environment, and of course, the table tennis.

Another institution that I am very thankful to, is the Science for Life Laboratory. I appreciate the sequencing services and bioinformatic support that I have been extensively using during my PhD. I have received help to advance my bioinformatic skills through various courses, seminars, and especially through the Swedish Bioinformatics Advisory Program. Estelle, thank you so
much for being there for me when I was making the first steps in the world of bioinformatics.

I also want to highlight the people who have been with me in the field. Spending a few weeks monitoring Swedish arctic foxes was a dream and a challenge at the same time, and I’m glad Henrietta and Bodil were there with me. Love and Anders, thank you for making me a part of the Dream Team, so that I could set off on the crazy adventures in the High Arctic. Åsa, Fred, Nico, and Tomas – I could hardly imagine a more inspiring group of people to spend field work with, and I am very thankful that you took me under your wings. Finally, the trip to Wrangel Island was the icing on the cake of this PhD. I’m glad that I could visit the refugium of the last mammoths with Love, Anders, Johannes, Maria, Alexei, and Gleb. Thanks to Gleb’s miraculous logistic and personal skills, everything worked out perfectly and I hope we’ll meet in the field again. Love, Gleb, and Igor, the days spent driving Argo and picking up tusks on the go will always belong to my favourite memories. I am also thankful to the Chaun Team – Peter, Susanna, Christine, Anders, Jesper, Kristaps, Harald, and my half-PhD-sibling Rasmus, it was great that we could spend those long days in Pevek together. Last but not least, I want to thank the Swedish Polar Research Secretariat for the support on these expeditions and the APECS Sweden team for the opportunities to get to know the polar community.

I want to acknowledge all my collaborators and everybody who contributed to my research. Above all, thanks to Sergey, Pavel, and Alexei for providing us with excellent samples, without which this thesis couldn’t happen.

This thesis was a group effort and I want to thank Rodrigo for drawing a stunningly beautiful cover, Johanna for translating the Swedish summary, David for reading my nailing page, and Love for commenting on the kappa.

I am also thankful to my former supervisor, Natália, for giving me the chance to do the kind of science I wanted to do. Natália, without your enthusiasm and pushing me out of my comfort zone, I wouldn’t get this far.

Finally, there are no words to express how much I am thankful for the possibility to do what I love and that is entirely thanks to my family, who always supported me in everything I have decided to do. Mami, tati, Laura – ďakujem vám za všetko, bez vás by som to nezvládla. I am glad that my family and friends are always there for me when I come back home. Thanks to Katka, Nion, Evoun, Terika, Lucik, Monila, NinaS, and Viki, for being friends for life.

Martin, thank you for selflessly supporting me, for being the worst critic and the biggest fan. You are my soulmate and I am incredibly thankful to have you on my side through everything that life brings along.