Ancient DNA from Sediments

and

Associated Remains

James Haile



Henry Wellcome Ancient Biomolecules Centre
Department of Zoology
University of Oxford
Trinity Term, 2008

James Seymour Haile
Pembroke College

Doctor of Philosophy
Trinity Term 2008

Ancient DNA from Sediments and Associated Remains

Abstract

This thesis explores the potential of new substrates for ancient DNA studies and addresses novel questions that can now be asked. It also highlights an additional use of ancient DNA extracted from a traditional source.

Acknowledgements

There are many people to whom I owe a debt of gratitude, and who each contributed in varying ways to making this thesis possible. Thanks must first go to my boss when I was Laboratory Manager of the Henry Wellcome Ancient Biomolecules Centre, Professor Alan Cooper, who offered me the chance to return to academia after a break of a few years. I would like to thank Dr. Beth Shapiro, who became my supervisor after inheriting me from Alan, and then to Dr. Cristian Capelli who inherited me from Beth!

My external supervisor, Professor Eske Willerslev has been a constant source of encouragement and enthusiasm throughout my Ph.D., as have numerous colleagues, friends and collaborators including Drs. Tom Gilbert, Greger Larson, Mike Bunce and Ross Barnett.

I am grateful to Malgosia Nowak-Kemp, Vertebrates Collection Manager of the Oxford Museum of Natural History, for her support, and especially to my wife Ruth, and children, Emily and Jack, for their love.

Finally I would like to dedicate this thesis to the memory of my father, Professor Neville Haile, who gave me a life-long interest in science, and who is himself the finest scientist I have known.

Contents

Abstract **2**

Acknowledgements **3**

Chapter 1 **7**
Introduction and Scope of Thesis **7**
History of ancient DNA 8
1.1.1 History of Sedimentary DNA 9
1.1.2 Sedimentary Ancient DNA 9
The entry of DNA into the soil 10
1.2.1 The fate of DNA in the soil 12
Problems in aDNA research 21
1.3.1 Sample-Based Contamination 21
1.3.2 Amplicon Contamination 22
1.3.3 Contamination controls 24
1.3.4 Jumping PCR 24
1.3.5 Abiotic DNA damage and base modification 25
Problems in *Sed*aDNA research 26
The Scope of this Thesis 28

Chapter 2 **32**
Materials and Methods **32**
2.1 Introduction 33
2.2 Sample Selection 33
2.3 The extraction of ancient DNA 33
2.3.1 The extraction of ancient DNA from sediments 34
2.3.2 Extraction of DNA from Eggshell 37
2.3.3 Extraction of DNA from Bone 38
2.4 Amplification of DNA by PCR 38
2.5 PCR Product Purification 40
2.6 Sequencing of DNA 40
2.7 Cloning of DNA 41
2.8 Amplification and sequencing of DNA by Roche Genome Sequencer FLX System (454) 42
2.9 Sequence Authentication 45

Chapter 3 **46**
Tracking Down the Last Mammoths: **46**
Proceed with Caution! **46**
3.1 Abstract 47
3.2 Introduction 47
3.3 Sites and Sample Information 48
3.4 Methods 50
3.4.1 Dating 50
3.4.2 DNA sequence authenticity 53
3.4.4 DNA extraction and amplification 55
3.4.5 Taxonomic identifications 57
3.5 Results 58
3.6 Statistical modelling of Taimyr macrofossil dates 62

3.7 Discussion 62

Chapter 4 70
DNA-Based Biodiversity Assessment: Plant DNA from Arctic Sediments 70
4.1 Abstract 71
4.2 Introduction 71
4.3 Sites and Sample Information 72
4.4 Methods 74
4.4.1 DNA Extraction and Amplification 74
4.4.2 Pollen Analysis 77
4.4.3 Dating 79
4.5 Results and Discussion 79
4.5.1 Botanically-Identified Plant Species Diversity 79
4.5.2 DNA-Identified Plant Species Diversity 79
4.5.3 Pollen-Identified Plant Species Diversity 81
4.5.4 Comparison of Plant Floral Survey, Plant DNA and Plant Pollen data 84
4.5.5 Ancient Sediments 88
4.6 Conclusions 95

Chapter 5 97
Megafaunal response to Catastrophic Change 97
5.1 Abstract 98
5.2 Introduction 99
5.3 Materials and Methods 101
5.4 Sampling, DNA extraction, and amplification. 105
5.5 Quantitative Real-Time PCR (qPCR) 108
5.6 Sequence Identification 109
5.7 Results and Discussion 111
5.7.1 Avian DNA 111
5.7.2 Plant DNA 119
5.7.3 Sheep DNA 119
5.8 Conclusion 122

Chapter 6 125
The Provenancing of Animal Remains from Archaeological Sites using Ancient DNA 125
6.1 Abstract 126
6.2 Introduction 126
6.3 Sites Descriptions and Locations 129
6.3.1 Mauritius 129
6.3.2 Lord Howe Island 131
6.4 Samples 135
6.4.1 Mauritius Samples 135
6.4.2 Lord Howe Island Samples 135
6.5 Methods 136
6.6 Results 141
6.6.1 Mauritius Dugong 141
6.6.2 Mauritius Pigs 141
6.6.3 Lord Howe Island Pigs 141
6.6.4 Mauritius Sediments 143

6.7 Discussion 146
6.7.1 Mauritius 146
6.7.2 Lord Howe Island 149
6.8 Conclusions 150

Chapter 7 **152**
Ancient DNA from Madagascan Elephant Bird Eggshells **152**
7.1 Abstract 153
7.2 Introduction 153
7.3 Sites and Sample Information 156
7.4 Methods 159
7.4.1 DNA Extraction and Amplification 159
7.4.2 Phylogenetic Analysis 159
7.5 Results 160
7.5.1 Eggshell 160
7.5.2 Sediments 161
7.6 Discussion 167
7.6.1 Inter-island Phylogeography and Phylogenetics 167
7.6.2 Intra-island Phylogeography and Phylogenetics 168
7.6.3 Ancient DNA in Eggshell 170
7.6.4 Ancient DNA from Sediments 172
7.7 Conclusions 173

Chapter 8 **174**
Conclusions and Future Directions **174**

Literature Cited **178**

Appendices **194**

Chapter 1

Introduction and Scope of Thesis

To see a World in a Grain of Sand
And a Heaven in a Wild Flower,
Hold Infinity in the palm of your hand
And Eternity in an hour.

William Blake, *Auguries of Innocence*

History of ancient DNA

The birth of the discipline of ancient DNA is widely agreed to have occurred in 1984 at Berkley when Russ Higuchi and colleagues reported that traces of DNA had been recovered from a museum specimen of the extinct Quagga, *Equus quagga* (Higuchi et al. 1984). Furthermore, the DNA extracted from the 150 year old remains was found to be of sufficient size and quality to allow the phylogenetic relationships of this species to be determined. Unsurprisingly, the results showed that the Quagga was a zebra, but the principle that was established was to prove to be more far-reaching.

The Quagga phylogenetic findings were based upon DNA painstakingly extracted and then sequenced through bacteria cloning, and although these methods were replicated by Svante Pääbo (1986) in museum specimens and thousands of years old human mummies it was the development of the Polymerase Chain Reaction (PCR) in the late 1980s (Mullis and Faloona 1987; Saiki et al. 1988) that led to the rapid expansion of the field.

Unfortunately with this rapid expansion came unsubstantiated claims for the amplification of DNA from multi-million year old samples, typically from organisms preserved in amber; weevils (Cano et al. 1993), termites (DeSalle et al. 1992; DeSalle et al. 1993), bees (Cano et al. 1992), but also from Miocene plants (Poinar et al. 1993). These claims were popularised in the book and the film, 'Jurassic Park' (Crichton 1991). Then in 1994 and 1995, scientists claimed to have recovered dinosaur DNA from bones and eggs (Woodward et al. 1994; An C-C et al. 1995). Whilst these and other multi-million year old claims such as the 250 mya halobacterial sequences from halite (Fish et al. 2002) have now been refuted (Lindahl

1993), the recent publications purporting the retrieval of *Tyrannosaurus* amino acid sequences (Asara 2007; Schweitzer et al. 2007) may reawaken the possibility of the recovery of truly ancient biomolecules. However, an untempered desire to find molecules of life in ever older samples, such as the Cretan Mammoths (Poulakakis et al. 2006), continues to undermine the validity of less controversial findings.

1.1.1 History of Sedimentary DNA

Sedimentary DNA (*sed*DNA) research as distinct from sedimentary *ancient* DNA has its origins in work by Goring and Bartholomew (Goring and Barthohlomew 1952) who showed that nucleic acids formed complexes with clay minerals and that this greatly increased their resistance to microbial degradation. These initial findings were developed by Greaves and Wilson who were interested in the formation of organo-mineral complexes and their influence on the metabolism of organic matter in soil (Greaves and Wilson 1969; Greaves and Wilson 1970).

However, the study of *sed*DNA accelerated in the wake of concerns over genetically modified organisms (GMOs), and their release into the environment (Trevors 1996). In 1994, it was shown that extra-cellular DNA can persist in the environment and over short time-spans, at least, retain the ability to transform cells (Lorenz and Wackernagel 1994).

1.1.2 Sedimentary Ancient DNA

By contrast to the controversial and now fairly long history of ancient DNA research, that of sedimentary ancient DNA (*sed*aDNA) research has been much shorter and so far less contentious. The founding paper was that of Willerslev and colleagues (2003) which shows not only that it was possible to recover ancient DNA from permafrost

and temperate sediments *in the absence of macrofossils*, but also that one could detect changes in the composition of plant groups over time. This dramatic finding has surprisingly not resulted in a plethora of publications on this area of research, perhaps because of the difficulty of obtaining soil and sediment samples from the field compared to sampling bone and tissue remains from museum collections.

The entry of DNA into the soil

DNA is found in the cells of all living organisms, as well as in viruses. Upon an organism's death, DNA is degraded by intracellular nucleases (and other reactive chemicals), and the nucleases in turn may be degraded by intracellular proteinases. This degradation would inhibit nuclease activity (Chiter et al. 2000), and thus promote nucleic acid survival. However, as these are ATP-dependent reactions, after cell-death and exhaustion of ATP, the rate of enzymatic reactions rapidly slow down (Pietramellara et al. 2006), and usually most of the DNA is released into the environment, where it is an ubiquitous molecule, accounting for up to 10% of extractable P in soil (Trevors 1996; Turner and Newman 2005). Since organelles and plastids are to some extent protected from cytoplasmic enzymes by their organelle membranes, the potential for their release into the environment may be enhanced.

Cells containing DNA can also be shed from an organism during its lifetime, through skin flakes (Bunce et al. 2005), urine (Valiere and Taberlet 2000), faeces (Poinar et al. 1998; Hofreiter et al. 2000), eggshells (Strausberger and Ashley 2001), hair (Higuchi et al. 1988; Taberlet et al. 1993; Gilbert et al. 2004; Gilbert et al. 2007), feathers (Taberlet and Bouvet 1991), leaves (Poté et al. 2005), pollen (Meier and Wackernagel

2003), root cap cells (del Pozo and Lam 1998), or in living prokaryotes through the secretion of plasmid and chromosomal DNA (Lorenz and Wackernagel 1991).

When DNA is released, it is rarely as a pure molecule, but tends to be coated with polysaccharides, lipids and proteins, and is released with RNA, cellular debris such as organelles and pieces of cell wall. It should be noted that when bulk DNA is extracted from soils and sediments, the nature of the source of the DNA, (whether it is cellular or extracellular, contained within a biological material such as a bone or insect chitin fragment or bound to soil particles, or free), is not known.

In the case of plant *sed*DNA, which has been intensively studied due to the interest in GMOs, the bulk of the DNA entering the soil comes from sloughed off root cap cells (Hawes 1990; de Vries et al. 2003). The majority of this DNA is cleaved into small fragments by endonucleases before it reaches the rhizosphere (Wang et al. 1996). The contribution of nuclear DNA from pollen to soil DNA depends upon the time of year, and species of plant, but can be considerable, as has been shown for sugar beet, *Beta vulgaris* (Arumuganathan and Earle 1991). Leaf-litter contribution is again temporally limited, and the action of microbial enzymes facilitates the release of DNA into the rhizosphere. However, within fresh plant tissues, DNA is largely degraded within five days of addition to soil at 21.5 °C (Poté et al. 2005). Finally pathogen colonisation can also lead to DNA being released from plant cells, as occurs when pepper, *Capsicum annum* is infected by *Xanthonomonas campsetris* pv. *vesicatoria*, and unlike in the case of programmed cell death, this infection results in the release of unfragmented DNA into the rhizosphere (Polverari et al. 2000).

1.2.1 The fate of DNA in the soil

Upon entry into the soil, the extracellular DNA molecule has three potential fates (Figure 1.1):

i. the incorporation into a bacterial genome through natural transformation,
ii. metabolism by bacterial and fungal exonucleases,
iii. persistence through binding to the soil matrix.

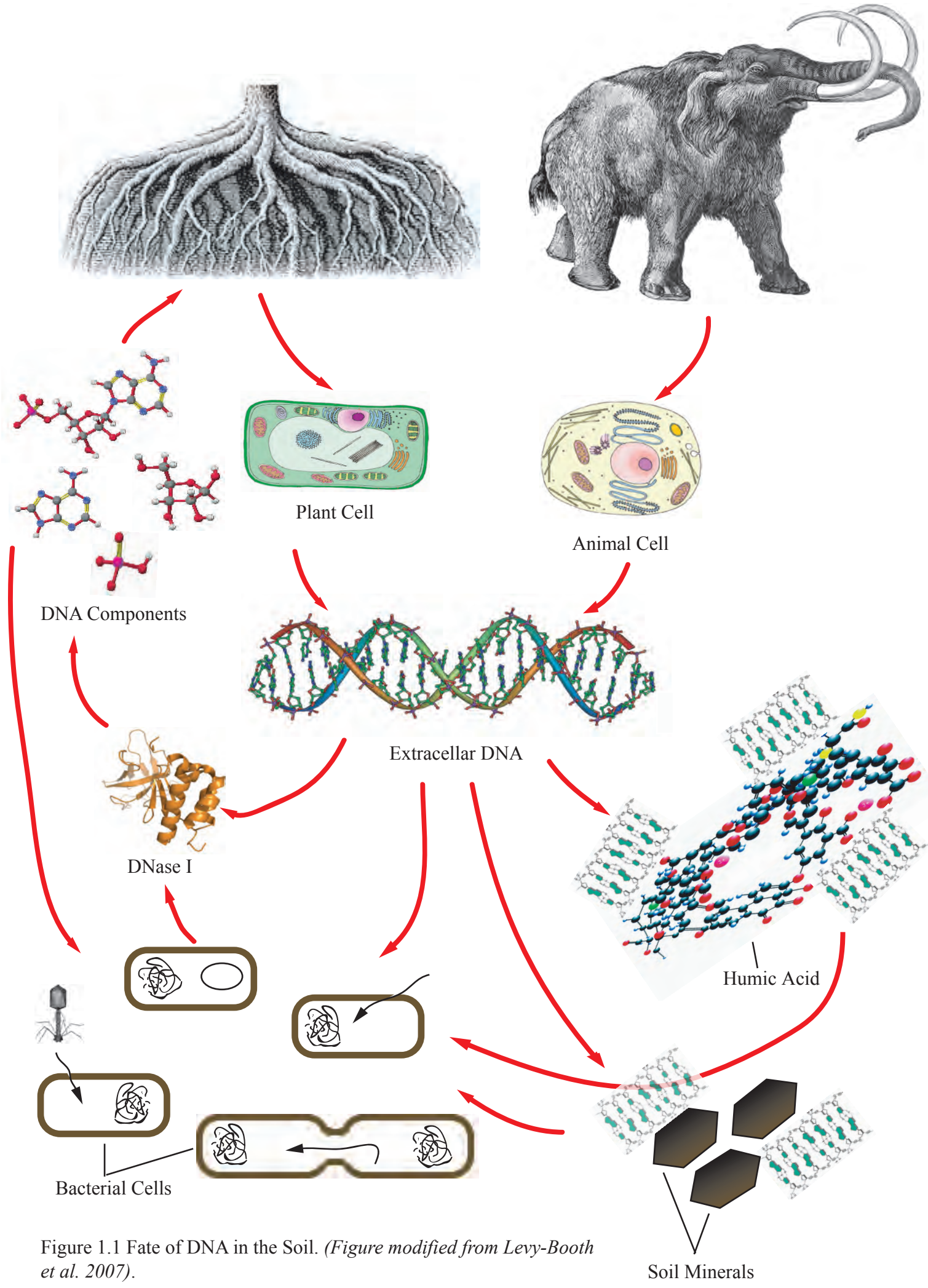


Figure 1.1 Fate of DNA in the Soil. *(Figure modified from Levy-Booth et al. 2007)*.

1.2.1.1 Natural Transformation

About 90 species of prokaryotes have been shown to be capable of natural or genetic transformation. These species include the soil bacteria *Acinetobacter* spp. and *Bacillus subtilis* (de Vries et al. 2003).

Bacterial transformation requires the release of extracellular (or 'naked') DNA into the environment, its persistence, and the development of competence by potential recipients, followed by the uptake, replication and expression of the DNA (Crecchio and Stotzky 1998). If these criteria are met, then it is possible that ancient DNA that has been integrated into the bacterial genome, may spread through the lateral gene transfer network, and so persist for some time at least in a living organism. Once incorporated into the bacterial genome, mutations may occur. If bacterially-incorporated aDNA were then to be extracted from sediments, base changes that occurred after the death of the original host would be encountered. This change is distinct from *post-mortem* damage, which occurs outside a living organism (Gilbert *et al.* 2003).

The extent to which bacterial transformation of aDNA occurs is unknown, but it is not likely to be significant. Bacteria are aquatic organisms, and the transfer of genetic material is highest when the soil water tension is near or at the optimum value for their growth (Stotzky et al. 1989). However, when there is sufficient water for enzymatic digestion of DNA, extracellular free DNA in soil is rapidly digested by soil nuclease activity (Trevors 1996). When the DNA is bound to Ca-montmorillonite clays, it can retain transformational activity for up to 15 days (Gallori et al. 1994). In

dry and/or frozen soils, with little free water, enzymatic digestion of DNA is expected to be very slow to negligible, but so too is transformational frequency.

Furthermore, successful transformation necessitates the presence of competent cells, and in a bacterial population this ranges from 0-100%. Competence is caused by a number of factors including exponential growth, post-exponential growth stasis, the concentration of soluble competence factors and high population density (Levy-Booth et al. 2007), yet within the soil, most bacteria live in a state of dormancy (Nielsen *et al.* 1997; Nielsen *et al.* 1997), or are limited in their ability to enter exponential growth phase (Bertolla et al. 1997). It has alternatively been suggested that competence is not primarily related to bacterial density, but to nutrient availability, and hotspots (or microbialspheres) containing high levels of nutrients might naturally occur in: the detritusphere, around the remains of a decaying organism; the plant rhizosphere (the soil surrounding plant roots) (Lorenz and Wackernagel 1991); the drilosphere (part of the soil influenced by earthworm secretions and castings); the porosphere (soil pores frequently occupied by young roots); preferential flow pathways that facilitate the movement of water and solutes through the soil (Flury et al. 1994), and caused by voles, earthworms, cracks *etc.*; and in biofilms, which are complex aggregations of microorganisms marked by the excretion of a protective and adhesive matrix.

Once the foreign fragment of DNA has been taken up by the cell, it then needs to become incorporated into the bacterial genome. Foreign DNA that remains in the cytoplasm is quickly degraded by nucleases and enters the internal DNA metabolism cycle. Regions of high similarity of between 25 and 200 bases are required between

the foreign and chromosomal DNA and so for eukaryotic DNA uptake transformation of mitochondrial DNA would be expected to be favoured over nuclear DNA.

Even for the most promiscuous and least discerning bacteria successful transformation of homologous regions only occurs between 0.1% in *Acinetobacter baylyi* to 25-50% in *S. pneumoniae* and *B. subtilis* (Palmen and Hellingwerf 1997; Thomas and Nielsen 2005). For non-homologous integration, the frequency of occurrence in *Acinetobacter* sp. was at least 10^9- fold lower than for homologous DNA (de Vries et al. 2003).

1.2.1.2 Metabolism by bacterial and fungal exonucleases

Soil bacteria can either synthesise DNA *de novo* or partially breakdown soil DNA and salvage the breakdown products. Due to the relatively high energy requirements of *de novo* synthesis it is thought that the latter pathway predominates (Levy-Booth et al. 2007) and bacterial DNases are believed to be the primary mechanism for extracellular DNA degradation in soil (Blum *et al.* 1997).

Bacteria actively secrete restriction endonuclease I into soil and this cleaves the DNA into smaller duplex oligonucleotide fragments of about 400 bp in length (Blum *et al.* 1997). *In vivo* experiments have shown complete restriction to about 7 bp fragments (Mishra 2002) and even hydrolysis to single nucleotides through the hydrolytic cleavage of phosphodiester linkages in the DNA backbone (Kornberg 1974), but this has not been demonstrated to occur in soils.

The amount of DNA in soils is often the limiting factor in bacterial growth. Twelve hours after 50 μg of DNA provision, Blum and colleagues (1997) found that the

number of viable microorganisms increased by one order of magnitude, and that 68% of the DNA added was enzymatically degraded to small molecular size. The rate of DNA sequence degradation is related to DNase concentration, temperature and the interaction of soil particles (see next section). Free DNA is completely degraded by DNase I concentrations above 1 μml^{-1} (Demanèche et al. 2001), and the half-life of extracellular DNA in leachate water decreased with increasing temperature (Gulden et al. 2005), supporting the hypothesis that this is a enzymatic reaction. Although at sub-zero temperatures nuclease activity is reduced, it has been shown that microbial DNA degradation can occur in temperatures of –8.4 °C (Diaz-Ravina et al. 1994). Nevertheless, moa (Aves; Dinornithiformes) DNA from 3000 year old dry temperate sediments has been recovered (Haile et al. 2007), mammoth DNA dating to 30 000 BP from permafrost sediments has been amplified, as well as 40 000 year old plant DNA (Willerslev et al. 2003) and from dirty ice beneath glaciers DNA has been shown to survive for half a million years (Willerslev et al. 2007).

The effect of heat on DNA survival has been modelled by Smith and colleagues (2003), and although the study focuses on DNA in bones, it too shows a negative correlation between temperature (couched in terms of thermal age) and DNA survival. Whereas the short-term persistence of DNA in bones is thought to be governed by microbial action, over the longer term chemical degradation comes into play. This involves DNA destruction through condensation (Pischetsrieder et al. 1999) and hydrolysis of the purine bases (Lindahl and Nyberg 1972). Whilst these degradation pathways have not been explicitly examined for sediments, it is likely that they would also occur.

The deoxyribonucleotides produced by enzymatic DNA restriction may be hydrolyzed to form nucleosides by bacterial catabolic enzymes excreted into the soil matrix. Nucleosides are more soluble than nucleotides, and can be readily transported into bacterial cells, and can then be further cleaved into their base (purine or pyrimidine) plus deoxyribose. Bases can be metabolized to uric acid, PO_4^{3-}, CO_2, NH_3 and enter their respective nutrient cycles, or be re-incorporated into nucleotides and enter the DNA salvage pathway.

1.2.1.3 Persistence through binding to the soil matrix

Persistence of DNA in the soil over time is a prerequisite for *sed*aDNA studies. Its endurance is a function of the interplay between the physical, chemical and biological properties of the soil. DNA released into the soil binds to clay and sand, humic substances and organomineral complexes, and this adsorption protects it from extracellular, microbial DNases and nucleases, which readily degrade free DNA in the soil solution (Greaves and Wilson 1969; Blum *et al.* 1997; Crecchio and Stotzky 1998; Khanna *et al.* 1998).

There are two ways in which soil binding can aid DNA survival;

i. by adsorbing DNases and nucleases thereby inhibiting their ability to hydrolyze extracellular DNA (Blum *et al.* 1997), and
ii. by shielding adsorbed DNA from nuclease activity (Crecchio and Stotzky 1998).

The greater the surface area and cation exchange capacity of soil particles, the stronger the affinity to extracellular DNA. The clay mineral montmorillonite (2:1 Si:Al) can absorb more than its own weight in DNA (Greaves and Wilson 1969) and is more than the clay mineral kaolinite is able to absorb (Khanna and Stotzky 1992; Pietramellara et al. 2001). Humic acids (HA) are the product of enzymatic degradation of numerous molecules including lipids, amino acids, peptides and polysaccharides. HA are resistant to decay, have a high degree of spatial heterogeneity, and like clays, are negatively charged, binding to DNA by means of cation bridging. Sand is less effective at binding DNA than clays due to its smaller surface area compared to clay minerals. However, adsorption to sand increases with increasing cation concentrations particularly divalent cations such as Ca^{2+} and Mg^{2+} which are more effective at forming sand-DNA bridges than monovalent cations (*e.g.* K^{+} and Na^{+}) (Figure 1.2) (Lorenz and Wackernagel 1987).

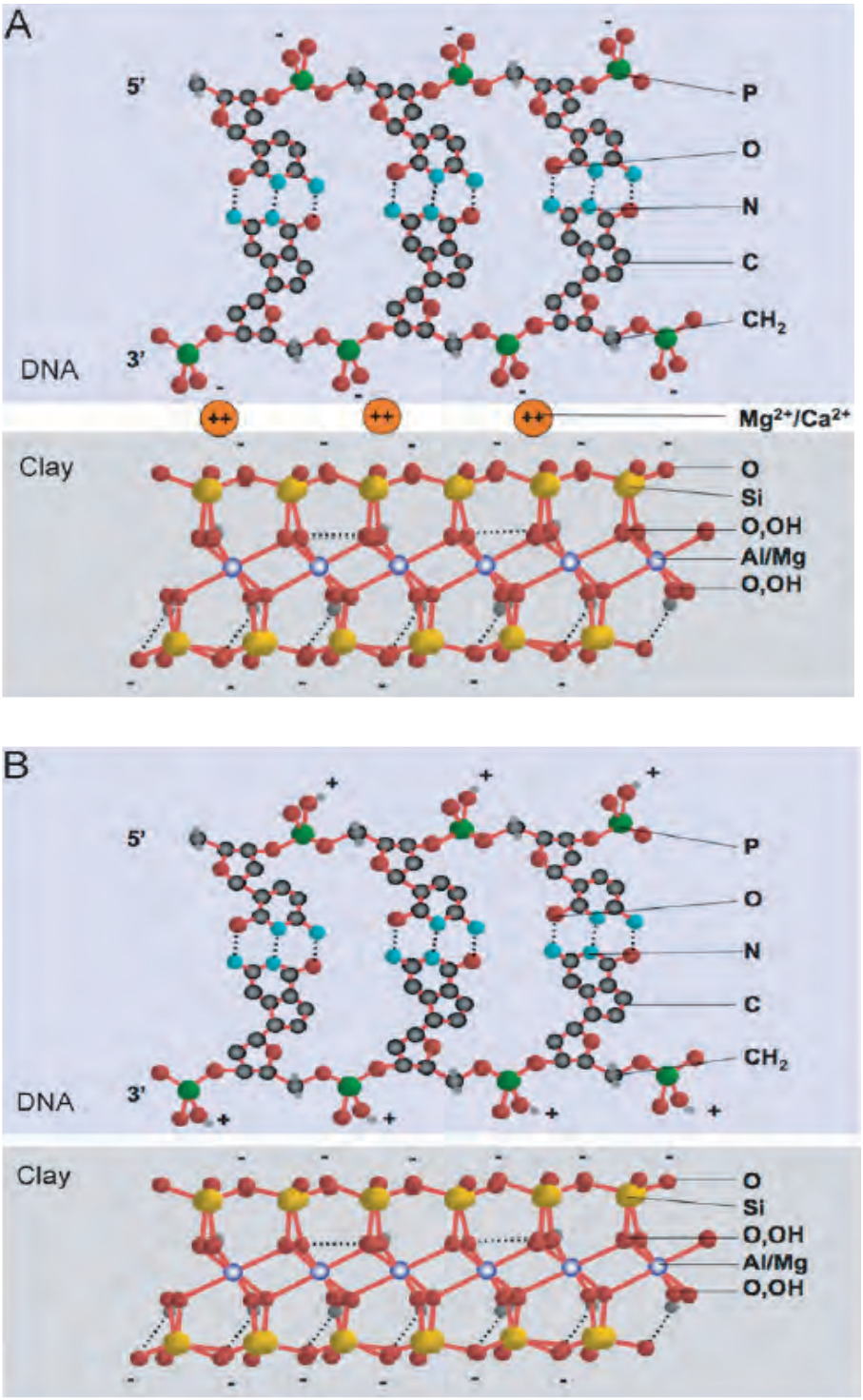


Figure 1.2 (A) Binding of Extracellular DNA to Clay at pH >5, mediated by cations. (B) Binding of Extracellular DNA to Clay at pH<5, occuring without mediation by Cations. *(Figure after Levy-Booth et al. 2007)*.

The effect of cations can be explained by the fact that above pH 5 DNA has a net negative charge as do the soil particles (clays, sand and HA), thus requiring positively charged ions to enable binding. Below pH 5, DNA has a net positive charge, and the need for mediating cations is therefore removed (Greaves and Wilson 1969).

It is important to note that the referenced studies have been carried out on purified DNA and single-component soil particles, and of course natural soils are spatially and chemically heterogeneous mixtures, and far more complex. This means that in the field it is very difficult to predict *sed*DNA survival. Recombinant DNA has been shown to persist for between 77 days (Widmer et al. 1997) and at least 2 years (Gebhard and Smalla 1999), and DNA fragments can survive much longer, especially in permafrost or in conditions where free water is absent (Willerslev et al. 2003; Haile et al. 2007; Willerslev et al. 2007).

Problems in aDNA research

1.3.1 Sample-Based Contamination

Of all the problems associated with aDNA research, by far the most significant is that of contamination. Sample-based contamination by exogenous DNA can occur at any stage following the death of an organism, and can include pre-extraction and post-extraction sample-sample cross-contamination, contamination with modern DNA, and contamination with PCR products, or amplicons. Although any of these routes of contamination can occur when dealing with modern samples, they are of particular concern in the field of ancient DNA because in any ancient sample the copy number of the aDNA is likely to be very low, and consequently the amount of endogenous aDNA is likely to be grossly outweighed by contaminating DNA especially if it is modern. The relatively few intact DNA molecules present in ancient samples explains

why the early attempts at bacterial cloning were so inefficient (Higuchi et al. 1984; Higuchi et al. 1987; Pääbo et al. 1989; Noonan et al. 2005). However PCR brought with it its own set of problems, for although under ideal conditions it enabled the repeatable recovery of much smaller amounts of DNA, it also allowed the possibility of erroneous results deriving from the amplification of non-target DNA, nuclear mitochondrial copies of DNA (numtDNA), or chimaeric products via jumping PCR.

1.3.2 Amplicon Contamination

In addition to sample-based contaminants are laboratory-sourced contaminants, and there are two types of these: PCR products (also know as amplicons), and contamination of laboratory consumables and plastic ware by modern DNA. Amplicons are ubiquitous in molecular biology labs where PCR takes place, are highly mobile, and numerous (1.1×10^{12} molecules generated per 25 μl PCR reaction). If these molecules are then extracted from ancient samples together with endogenous DNA, and the primer target region of the DNA is the same as that of the amplicons, then the potential to produce false positives is high (Kwok and Higuchi 1989), especially as the number of PCR cycles used in aDNA is typically large, up to 60.

In order to reduce the likelihood of amplicon contamination, ancient DNA laboratories are usually physically separated into ones where DNA extraction takes place, and ones where PCRs occur, with strict guidelines on equipment and people moving between the two. However, the problem of amplicon contamination could simply be removed by adding Uracil-DNA glycosylase to the PCR mix and replacing the base thymine in the PCR dNTP mix with the base uracil, with the resulting amplicons also possessing this base change. Before the first round of PCR, conditions

can be adjusted so that the Uracil-DNA glycosylase will catalyse the hydrolysis of the N-glycosylic bond between the uracil and sugar, leaving an apyrimidinic site in any contaminating uracil-containing single or double-stranded DNA created by a previous amplification (Lindahl 1977; Lindahl et al. 1977). As the polymerase would be blocked by the abasic site, the contaminating DNA would not be amplified thus controlling any amplicon, or carry-over contamination (Longo et al. 1990).

One disadvantage of using uracil-DNA glycosylase is that any genuine single-stranded aDNA that contains a single cytosine deamination to uracil would not be amplified, leading to a potential loss of genetic information from the undamaged proportion. If one is trying to recover information about past floral and faunal biodiversity from the sediments (rather than trying to carry out phylogenetic reconstructions), then single base errors are unlikely to lead to incorrect species identification. In addition, in sediments at least, cytosine deamination does not appear to be a major cause of DNA damage (*pers. obs.*).

1.3.3 Contamination controls

Several criteria for aDNA studies have been published, and the suggested measures for contamination control variously include the use of UV irradiation, bleach, disposable plasticware, physical separation of pre and post-PCR labs, independent replication, use of extraction blanks and negative PCR controls, and amino-acid racemisation (Handt et al. 1994; Poinar et al. 1996; Cooper and Poinar 2000; Poinar 2003; Shapiro et al. 2004). Gilbert and colleagues (2005) take a more rational and less proscriptive view of the criteria suggesting that even if all the precautions were taken, incorrect conclusions could still be reached, and conversely, defendable conclusions could also be achieved without the strict adherence to the guidelines. It also very much depends upon the organisms of interest; human and bacterial aDNA studies are much more likely to encounter false positives than studies of plants and animals that are distantly related to modern organisms, especially if they are from extinct species. Lying intermediate between human and bacteria studies, and studies of aDNA from extinct species, with regard to contamination risk, is investigations into aDNA from domestic species, as DNA from these animals can often occur as laboratory contaminants (Leonard et al. 2007).

1.3.4 Jumping PCR

Jumping PCR is caused during the PCR process, when the polymerase encounters a block to replication caused by, for example, covalent cross links (Pääbo et al. 1989; Mitchell et al. 2005) or abasic sites, and then stops, with the template and the partly extended primer then dissociating (Pääbo et al. 1990). During the next round of amplification the extended primer can bind to a different and undamaged template and copy the template until it reaches the complementary end primer. The sequence thus

formed is a chimaeric fusion of two different templates from two separate amplification cycles. If this chimaera is created early in the PCR cycle the products can come to dominate the real amplicons, and thus lead to erroneous results.

The formation of crosslinks in *sed*aDNA appears to be the leading cause of DNA damage in frozen sediments, occurring ~100 times more frequently than single- and double-strand breaks, and also more frequently than depurination (Hansen et al. 2006). Thus it would seem that amplification of *sed*aDNA should be particularly prone to jumping PCR, but this does not seem to be the case (*pers. obs.*).

1.3.5 Abiotic DNA damage and base modification

Post-mortem DNA damage by microbial action has been outline above. In addition, the DNA molecule can be broken by the action of water molecules which cleave the phosphate backbone of DNA (Lindahl 1993; Pääbo *et al.* 2004), resulting in the fragmentation of regions of contiguous DNA to much smaller fragments (typically 100-500 bp). Water can also lead to the deamination of bases and breakage of the glycosyl base-deoxyribose bond through hydrolysis, leading to modified bases and abasic sites respectively (Lindahl 1993; Pääbo *et al.* 2004).

A particularly common problem is the deamination of cytosine and methyl-cytosine which causes base pairing with the incorrect complementary base during PCR (Higuchi et al. 1987). More explicitly, the transitions caused are thymine to cytosine (T→C) and adenine to guanine (A→G) (Type I damage); and cytosine to thymine (C→T) and guanine to adenine (G→A) (Type II damage) (Hansen et al. 2001).

In vivo and *in vitro* studies have shown that under similar conditions, the rate of deamination of cytosine is over twenty times that of adenine (Lindahl 1979), resulting in the predominance of Type II damage. The presence of such modified nucleotides in samples with low copy number DNA can manifest themselves as sequence misidentification. Furthermore, the distribution of these damage sites within sequences appears to be non-random, occurring in naturally variable sites, or hotspots (Gilbert *et al.* 2003; Banerjee and Brown 2004; Gilbert *et al.* 2005).

Problems in *Sed*aDNA research

In addition to the problems of "traditional" aDNA research outline above, one further problem that is peculiar to *sed*aDNA studies (although shared with other archaeological and palaeontological pursuits), is that of temporal fidelity. When aDNA is extracted from macrofossil remains, it is usually possible to date the remains directly by AMS radiocarbon dating or by some other means. However, in the case of *sed*aDNA, when the DNA is extracellular, one has to rely upon its association with datable remains. Radiocarbon dating can be carried out on macrofossils extracted from the sediments, and from bulk sediments (McGeehin et al. 2001; DeLong and Arnold 2007), and by optically stimulated luminescence of single-grain quartz crystals (Aitken 1998; Lian and Roberts 2006; DeLong and Arnold 2007; Arnold et al. 2008; Demuro et al. in press). These researchers have shown that in permafrost environments, at least, there does not appear to be significant bioturbation and reworking of sediments, outside of obvious locations such as lacustrine and river systems, and in the vicinity of ice wedges. Unfortunately, DNA molecules cannot be directly dated and subsequently amplified, and even if sediments are chronologically consistent, it remains a possibility that *sed*aDNA recovered from them is intrusive.

This vertical movement of DNA is thought to be more of a problem in temperate locations, or at least those sites where there is movement of water through the soil profile. The movement of the DNA molecule has been demonstrated by both water leaching and water rise by capillarity (Poté et al. 2003; Agnelli et al. 2004; Asher et al. 2005), and in New Zealand, Haile et al. (Haile et al. 2007) showed that the co-occurrence of sheep DNA with moa DNA was due to DNA-containing sheep urine percolating down through the soil strata. By contrast, the moa DNA was found to be stratigraphically discrete because there was no free water in the sites, and because moa being birds do not excrete nitrogenous wastes as liquid urea (see Chapter 5).

Several lines of evidence suggest that vertical movement of *sed*aDNA is not a serious issue in permafrost environments.

i. Endurance of *sed*aDNA is reliant upon its binding to soil particles and humic acids and any DNA that becomes released is readily degraded (Levy-Booth et al. 2007).
ii. It appears difficult to release bound *sed*aDNA using water alone (see Chapter 3).
iii. Hansen and colleagues (2006) have shown that there is a clear relationship between the age of sediments that DNA has been extracted from, and the degree of molecular damage (single-strand breaks and crosslinks).
iv. The different taxonomic groupings of plant *sed*aDNA from permafrost sediments across time (Willerslev et al. 2003).

The Scope of this Thesis

Willerslev and colleagues (2003) in their seminal publication demonstrated that stratigraphically consistent mammal and plant ancient DNA could be recovered from permafrost sediments. The key objective of this thesis has been to take this finding, adapt and improve the methodology, and to test specific hypotheses that their potentially revolutionary discovery allows. I have also been keen to explore the limits of *sed*aDNA, both in terms of preservation at different latitudes, as well as preservation in sediments of different ages and to investigate the information yielded by targeting the DNA of different species.

Chapter 2 describes in detail the materials and methods used and developed for this thesis.

Chapters 3 – 7 form the core of the thesis, which adopts the 'paper approach'. With the exception of Chapter 5 which is modified from a published paper (Haile *et al.* 2007), the chapters are taken from papers which are in a pre-submission stage of the publication process.

In Chapter 3, *sed*aDNA of arctic megafauna recovered from dated permafrost sediments is used to look at faunal change over time, and to investigate the possibility of establishing more recent last occurrence dates for mammoth and other megafauna on the continents of America and Asia. This chapter also addresses the potential problem of reworking of permafrost sediments, and looks at different dating techniques that can shed light on this issue.

The samples used in this chapter were collected by Professor Eske Willerslev and Dr Tom Gilbert; the OSL dating was conducted by Dr Jean-Luc Schwenninger; radiocarbon dating was carried out by Dr Tom Higham and Dr Fiona Brock; Dr Barry Brock did the statistical modelling of macrofossil dates; and sequences were analysed by Professor Rasmus Nielsen. I wrote the chapter, generated the data, and interpreted the results.

Chapter 4 also deals with DNA from permafrost, but focuses instead on modern and ancient *plant* DNA. By looking at the relationship between the distribution of living plants, the pollen identified from the topsoil, and the DNA extracted from the same soils, the aim of this study is to lay the foundations for future work that will take the findings and apply them to investigate changes in plant species composition over space and time as a proxy for interpreting past climatic fluctuation on the micro- and macro-scale. Several Pleistocene permafrost samples are investigated, and the results interpreted in terms of climate change. This study makes pioneering use of the 454 Life Sciences Genome Sequencer FLX System, which enables the sequencing of over 400 000 reads per run, thus drastically reducing the risk of missing low frequency sequences.

The samples used in this chapter were collected by Dr Kari Anne Bråthen who also carried out the plant survey; pollen analysis was conducted by Dr Henrik von Stedingk; and DNA was sequenced and analysed by Professor Pierre Taberlet and Mr Ludovic Gielly, who both also created the comparative arctic species DNA database. I wrote the chapter, extracted DNA from all of the samples, and interpreted the results.

In Chapter 5 the focus moves south to New Zealand, and exchanges plants for birds. In this chapter the response of different genera of moa to catastrophic environmental change is reported, and the vertical movement of DNA is addressed using quantitative PCR.

The samples used in this chapter were collected by Professor Richard Holdaway; Dr Tom Gilbert conducted the qPCR assays; and sequences were analysed by Professor Rasmus Nielsen. I wrote the chapter, extracted DNA from all of the samples, and interpreted the results.

The final two primary data chapters strive to achieve the hitherto unobtainable; sedimentary aDNA from the tropics, an environment that has not been productive for aDNA studies using conventional materials. The *sed*aDNA results from the locations listed are at best equivocal, and as a consequence it was decided to concentrate on associated macro-remains in addition to the sediments.

Chapter 6 describes a potentially new use of aDNA, one that is especially suited to the relatively simple ecosystems of islands, and that would be of particular interest to archaeologists: that of aDNA provenancing. In this chapter it is shown that it is possible to genetically determine the geographic origin for certain species that show a strong phylogeographic signal. This has implications for anthropogenically modified environments, human trade routes, economic systems etc.

The samples used in this chapter were collected by Dr Kimberley Owens, Professor David Blair and by myself; Dr Greger Larson extracted DNA from the Lord Howe Island samples, and Dr Kimberley Owens interpreted the results from that location;

Dr Beth Shapiro conducted the phylogenetic analysis. I conceptualised the project, did the remaining lab work, wrote the chapter, and interpreted the Mauritius results with the help of Dr Ranjith Jayasena, Dr Kenneth Rijsdijk and Dr Julian Hume.

Chapter 7 evolved out of an attempt to extract aDNA from tropical sediments that were thought to be associated with elephant bird nesting sites and to contain fragments of eggshell. Whilst the sediment results are ambiguous, the refocusing of efforts onto macro-pieces of elephant bird eggshell has yielded significant findings as to their phylogenetic position, and lays the foundation for further investigations.

The samples were collected by Dr Jean-Luc Schwenninger; radiocarbon dating was carried out by Dr Tom Higham; and phylogenetic analysis was done by Dr Matt Phillips. I conceptualised the project, wrote the chapter, generated the DNA data, and interpreted the results.

Conclusions and implications for further work are discussed in Chapter 8.

Finally contained within the Appendices are tables relating to Chapter 4, and a list of co-authored papers.

Chapter 2

Materials and Methods

The men of experiment are like the ant, they only collect and use; the reasoners resemble spiders, who make cobwebs out of their own substance. But the bee takes the middle course: it gathers its material from the flowers of the garden and field, but transforms and digests it by a power of its own. Not unlike this is the true business of philosophy; for it neither relies solely or chiefly on the powers of the mind, nor does it take the matter which it gathers from natural history and mechanical experiments and lay up in the memory whole, as it finds it, but lays it up in the understanding altered and digested. Therefore, from a closer and purer league between these two faculties, the experimental and the rational (such as has never been made), much may be hoped.

Francis Bacon, *Novum Organum*

2.1 Introduction

This chapter describes the methods used to extract, amplify and sequence aDNA from a variety of samples including sediments, eggshell and bone.

2.2 Sample Selection

Sediment samples were mostly obtained by taking horizontal cores into river-cuttings, into the side of archaeological sections, or for modern material by taking plugs of sediment from immediately below the topsoil. The drilling core head was frequently spiked with a solution containing laboratory strains of *Serratia marcescens* bacteria, and in the laboratory PCRs to amplify this contaminant were carried out to ensure that the outer layer of the core had been successfully removed and discarded prior to DNA extraction (Willerslev *et al*. 2003).

Bones were collected immediately adjacent to sediment collections, as were eggshells where possible. Each sample was individually bagged and labelled. Appropriate methods were used when sampling, including using disposable tools, wearing gloves and facemask, cleaning with bleach etc.

2.3 The extraction of ancient DNA

Although the extraction of *ancient* DNA from samples is the objective, in reality *all* DNA is extracted indiscriminately, and selection starts only at the PCR stage, when due to the presence of contaminants non-target DNA may still be amplified.

2.3.1 The extraction of ancient DNA from sediments

Two very similar methods for ancient DNA extraction from sediments have been employed, which essentially differ only in the volume of sediment processed.

2.3.1.1 Extraction of sedimentary DNA using the FASTprep method

Approximately 1 g of wet weight sample is divided into two sub-samples of 0.5 g and placed in 2 ml FASTprep tubes (Qbiogene Inc., Carlsbad, CA, USA) containing 250 mg glass beads (106 microns and finer, Sigma-Aldrich, Steinheim, Germany). 600 µl of lysis buffer (0.02 g/ml Sarcosyl, 50 mM Tris-HCl, 20mM NaCl, 3.5% 2-mercaptoethanol, 50 mM 1,4-Dithio-L-threitol (DTT), 2 mM N-phenacylthiazone bromide (PTB), 0.8 g/ml Proteinase K) (Bulat et al. 2000) is then added.

(In the digestion buffer, Proteinase K is an endolytic serine protease which cleaves after aliphatic, aromatic and hydrophobic amino acids to abolish tertiary protein structure (Voet and Voet 1995). DTT reduces cystine crosslinks in proteins to destroy their quaternary structure and allow further degradation (Voet and Voet 1995). Sodium dodecyl sulfate (SDS) is a detergent which acts to denature proteins (*e.g.* nucleases) through interfering with noncovalent subunit interactions as well as solubilising biological membranes (Voet and Voet 1995). PTB cleaves glucose-derived protein crosslinks (Vasan et al. 1996) and has been empirically shown to increase the success of PCR reactions from ancient material (Poinar et al. 1998), although the exact mechanism by which it achieves this is unclear.)

The tubes are then placed in a FastPrep® Instrument (Qbiogene) where they are shaken for 45 s which causes samples to be pulverised, and cells lysed. Samples are then placed on ice for 2 min. Shaking and cooling are then repeated three more times, before the samples are placed on a rotator and left at 65 °C overnight.

The samples are then adjusted to 1.15 M NaCl and 300 µl of chloroform/octanol (24:1) added. The mixture is rotated for 10 min and after centrifugation at 12 000 g for 2 min, the supernatant purified using silica spin columns and PB-buffer (Qiagen DNA purification kit II), followed by washes in 0.5 ml Salton wash 1 and 2 (BIO 101), and 0.5 ml AW1 and AW2 (Qiagen tissue kit). These four wash steps are essential in order to remove co-extracted inhibitors such as humic acids. The DNA is then eluted twice with 100 µl EB buffer (Qiagen purification kit II) and stored at –20 °C.

2.3.1.2 Extraction of sedimentary DNA using the PowerMax Soil™ Kit

The major advantage using the PowerMaxSoil™ DNA Isolation Kit (Cambio) is that it allows the simultaneous processing of far greater volumes of sediment; typically up to 10 g wet weight of soil is processed in a 50 ml tube containing garnet grit, to which is added 12 ml of lysis buffer (Bulat et al. 2000) or C2, bead solution (guanidine thiocyanate solution) and C1 solution (sodium dodecyl sulfate solution). The sample is then vortexed for 5 min at the highest speed to ensure cell lysis and/or release of DNA from soil particles, and rotated overnight at 65 °C.

After centrifugation at 2500 × *g* for 3 min and the supernatant transferred to a clean tube containing 5ml of acetate solution and left to incubate at 4 °C for 10 min. The acetate solution acts to precipitate non-DNA organic and inorganic material, including

humic substances, cell debris, and proteins. It is important to remove contaminating organic and inorganic matter that may reduce DNA purity and inhibit downstream DNA applications.

The tube is centrifuged at 2500 x *g* for 4 min and the supernatant transferred to a clean tube containing 4 ml C3 solution and again left to incubate at 4 °C for 10 min. Solution C3 is a second reagent to precipitate additional non-DNA organic and inorganic material including humic acid, cell debris, and proteins.

After centrifuging at 2500 x g for 4 min, the supernatant is removed to a clean tube and 30 ml of solution C4 (guanidine HCl – isopropanol solution) added. This mixture is then spun through silicon spin filters. Solution C4 is a high concentration salt solution. Since DNA binds tightly to silica at high salt concentrations, this will adjust the DNA solution salt concentration to enable binding of DNA, but not non-DNA organic and inorganic material that may still be present at low levels, to the Spin Filters. 10 ml of solution C5, an ethanol based wash solution, is then spun through, and is used to further clean the DNA that is bound to the silica filter membrane in the Spin Filter. This wash solution removes residual salt, humic acid, and other contaminants while allowing the DNA to stay bound to the silica membrane.

Finally 5 ml of solution C6 (Tris buffer solution) is added to the spin filter membrane and centrifuged at 2500 x *g* in order to elute the bound DNA, and the elute is further concentrated to approximately 200 μl using a 30 000 MWCO Millipore Amicon ultracentrifuge tube. The DNA containing solution is stored at –20 °C.

2.3.2 Extraction of DNA from Eggshell

Approximately 1-2 g of eggshell was powdered using a Braun Mikrodismembrator U (B.Braun Biotech International, Germany) in a sterile steel canister using tungsten or stainless steel ball bearings and then the powder transferred to a 15 ml (or 50 ml) tube (Eppendorf, U.S.A) containing 12 ml of extraction buffer. The extraction buffer consists of Tris-HCl (20 mM pH8.0), sodium dodecyl sulphate (SDS, 1%w/v), dithiothreitol (DTT, 10 mM), proteinase K (2% w/v), and 0.5 M ethylene diaminetetraacetic acid (EDTA, Sigma, UK). Samples were then rotated for 48 h at 55 °C.

After 48 h, the samples are centrifuged (5 min at 6000rpm) and the supernatant decanted into an Amicon Ultra-30 centrifugal filter (Millipore) in order to reduce the volume to approximately 1 ml. 5 ml of PB Buffer (Qiagen) is added to the extract, which is then place in a QIAquick and the extract cleaned following the Qiagen QIAquick PCR purification protocol.

2.3.3 Extraction of DNA from Bone

The surface of bone samples was first cleaned of gross contaminants by abrading it using a Dremel 3981 professional drill fitted with a disposable emery wheel (No. 409). A subsample of up to 2 g is then taken, and powdered using a Braun Mikrodismembrator U (B.Braun Biotech International, Germany) in a sterile steel canister using tungsten or stainless steel ball bearings. The powdered samples are then placed in separate 15 ml tubes (Eppendorf, U.S.A) and 14 ml of 0.5 M pH8 ethylene diaminetetraacetic acid (EDTA, Sigma, UK) added to each, together with Proteinase K, and the samples rotated at 55 °C. After the sample has been digested, the tubes are centrifuged and the supernatant concentrated using Amicon Ultra-30 centrifugal filters (Millipore), and the extracts cleaned using QIAquick tubes, following the Qiagen QIAquick PCR purification protocol.

2.4 Amplification of DNA by PCR

PCR amplifications (25 µl total volume) used Platinum *Taq* High-Fidelity polymerase (1.25 U, Invitrogen, UK), 1x buffer (Invitrogen), bovine serum albumin (BSA, 1mg/ml, Sigma), $MgSO_4$ (2 mM) 2'-deoxynucleotide 5'-triphosphate mix (containing Adenosine, Guanosine, Cytidine, Deoxythymidine, all 25 mM), extract (1-5 µl), and forward and reverse primers (1 µM). The Platinum mix contains both *Taq* polymerase (from *Thermus aquaticus*) and *Pyrococcus* GB-D polymerase (which has 5' to 3' exonuclease activity), giving increased fidelity over *Taq* alone. The need for increased fidelity is warranted by the generally degraded nature of ancient DNA and tendency for mis-incorporations to occur during polymerisation (Hansen *et al.* 2001; Gilbert *et al.* 2003; Pääbo *et al.* 2004)□□.□ The mix also contains a thermostable antibody to Taq, preventing polymerase activity until thermal cycling has begun. BSA acts as a

non-specific chelator of potential inhibitors in the reaction mix, which are often co-extracted with ancient DNA samples.

Due to the individual nature of each DNA extract, it was frequently necessary to adjust the reaction conditions to ensure optimal PCR success. This often included altering the amount of extract added to the final reaction mix to increase template concentration (using >1 µl extract) or decreasing the amount of inhibition by serially diluting the extract. Adjustment of the final concentration of $MgSO_4$ or decreasing annealing temperature allowed for promiscuity of primer binding when the exact sequence of template bound by the primer was unknown i.e. when primers were designed from alignments of related species.

PCR was carried out using an MJ Research PTC 225 Tetrad thermal cycler (Bio-Rad, U.S.A) with individual cycling conditions given in subsequent chapters. Experiments also routinely incorporated negative controls (reactions set up as normal but with deionised water added instead of extract).

PCR products (5 µl) were electrophoresed on agarose gel (2%, Sigma-Aldrich, UK), stained with ethidium bromide and viewed under shortwave UV light (BioRad Gel Doc 2000). As a comparative standard ϕX174 DNA/*Hae*III ladder (Promega, UK) was also loaded alongside PCRs. Following successful amplification, the products were then purified for sequencing and cloning.

2.5 PCR Product Purification

Two separate PCR purification techniques were employed depending on the number of positive amplifications per reaction. If more than 10 reactions were successful, amplified products were purified using the Montage PCRμ96 system (Millipore, U.S.A) based on size-exclusion technology. PCRs were diluted with deionised water (80 μl) and transferred to a Montage plate, which was then vacuum filtered to dryness. Washing the membrane with deionised water (50 μl) removed primers, salts and deoxynucleotides before resuspending in more deionised water (30 μl).

If less than 10 reactions were successful, amplified products were purified using the Qiagen PCR purification Kit (Qiagen Ltd.). Using the supplied silica columns and buffers the amplified DNA was purified and eluted with deionised water (30 μl). All purified PCR samples were stored in the Department of Zoology at –20 ºC.

2.6 Sequencing of DNA

Direct sequencing of amplifications used ABI BigDye terminator v3.1 cycle sequencing kit (Applied Biosystems, U.S.A). Amplifications were directly sequenced in both directions. Total reaction (10 μl) consisted of BigDye mix (0.25 μl), sequencing buffer (1.875 μl), single primer (1.0 μl of 1mM solution), deionised water (5.875 μl) and purified amplification (1.0 μl). Thermal cycling of the resultant mix involved an initial heating step of 95 ºC for 1 min, followed by 25 cycles of 95 ºC for 10 s (strand separation), 50 ºC for 5 s (primer annealing), and 60 ºC for 2 min (polymerisation).

A final step at 20 ºC for 5 min, reduces the cycler to a more stable temperature for the sequencing product. To prepare the dye-labeled DNA for reading, ethanol

precipitation was used to remove unused dye chemistry, primers and salts. Molecular biology grade ethanol (64 µl of 95%), and deionised water (26 µl) were added to each reaction, which was then sealed, vortexed, centrifuged and left at room temperature (15 min). The mixture was then centrifuged (2750g, 4 ºC for 40 min) before inversion and further centrifugation (500 g for 1 min) to remove the ethanol. More ethanol (150 µl of 70%) was added and the mix was centrifuged again (2750 g, 4 ºC for 10 min) before a second inversion (500 g for 1 min). The precipitated pellet was then left for 10 min in the dark, at room temperature to ensure complete evaporation of ethanol before transferral to the DNA sequencing facility (Zoology Department, University of Oxford) for reading on an ABI Prism Capillary DNA Sequencer (377, 310 or 3700 model as appropriate).

2.7 Cloning of DNA

As well as direct sequencing, many PCR products were cloned using the TOPO TA Cloning kit (Invitrogen) with TOP10 or TOP10F' competent *Escherichia coli* cells. pCR 2.1-TOPO vector (containing amp^R gene) solution (0.25 µl) and salt solution (0.25 µl) was added to purified amplifications (1 µl), briefly microfuged and then incubated at room temperature for 5 min. Thawed, competent cells (20 µl) were then added to this and incubated on ice for another 5 min. The cell/vector/PCR mix was then heat-shocked (42 ºC for 30 s) and returned to the ice. SOC medium (80 µl) was added to aid cell recovery and growth and the mixture was incubated (38 ºC for 1 h) while being constantly agitated. Transformed cells were added to LB plates consisting of LB Agar (1.5% w/v), ampicillin (0.005% w/v), X-gal (4% w/v), and isopropyl-beta-D-thiogalactopyranoside (IPTG, 100 mM). The plates were then incubated overnight (38 ºC).

Identification of colonies containing the PCR insert is facilitated by the ability of the medium to select for cells that contain the pCR 2.1-TOPO plasmid (i.e. ampicillin resistant) and the disruption of the ß-galactosidase gene (i.e. unable to metabolise X-gal, which causes the absence of an accumulation of bright blue catabolites). White colonies therefore have both the plasmid (to enable growth) and the insert (to be unable to metabolise X-gal). Typically between 8-16 white colonies were selected with sterile pipette tips and transferred into deionised water (30 µl). This bacterial suspension was used to seed PCR amplifications using the primers M13R (5' GTC CTT TGT CGA TAC TG) and T7 (5' CCC TAT AGT GAG TCG TAT TA). Amplifications (12.5 µl total volume) used AmpliTaq Gold polymerase (0.75U, ABI), 1× buffer, 2'-deoxynucleotide 5'- triphosphate mix (all 25 mM), $MgCl_2$ (2 mM), bacterial suspension (5 µl) and M13R/T7 primers (1 µM each). Thermal cycling involved an initial activation step at 95 ºC for 2 min, followed by 40 cycles of 95 ºC for 45 s, 56 ºC for 45 s and 72 ºC for 1 min 30 s. Cloning PCR reactions were cleaned up using the Montage PCRµ96 system (Millipore, U.S.A). Bacterial clone solutions and purified PCR amplifications were stored in the Department of Zoology at –20 ºC.

2.8 Amplification and sequencing of DNA by Roche Genome Sequencer FLX System (454)

Amplification and sequencing of DNA using the Roche Genome Sequencer FLX System (454), is a cost-efficient way of generating huge numbers of sequences (>400 000) in a single reaction. Because it makes use of emulsion PCR (emPCR), it effectively allows numerous clones to be generated from PCR products. In conventional bacterial cloning one would typically pick 8-16 colonies (occasionally as many as 100). However, if using non-specific primers, the amplicons of contaminants,

or of amplicons that had been successfully amplified in the early stages of PCR would come to dominate the mix of PCR products, and hence possibly conceal presence of less-numerous amplicons.

The 454 system can sequence samples from a variety of starting materials, including genomic DNA, PCR products, BACs and cDNA. Because of the impurity of aDNA samples, PCR products resulting from non-specific primer amplification are normally used.

The operation of the 454 system is illustrated in Figure 2.1.

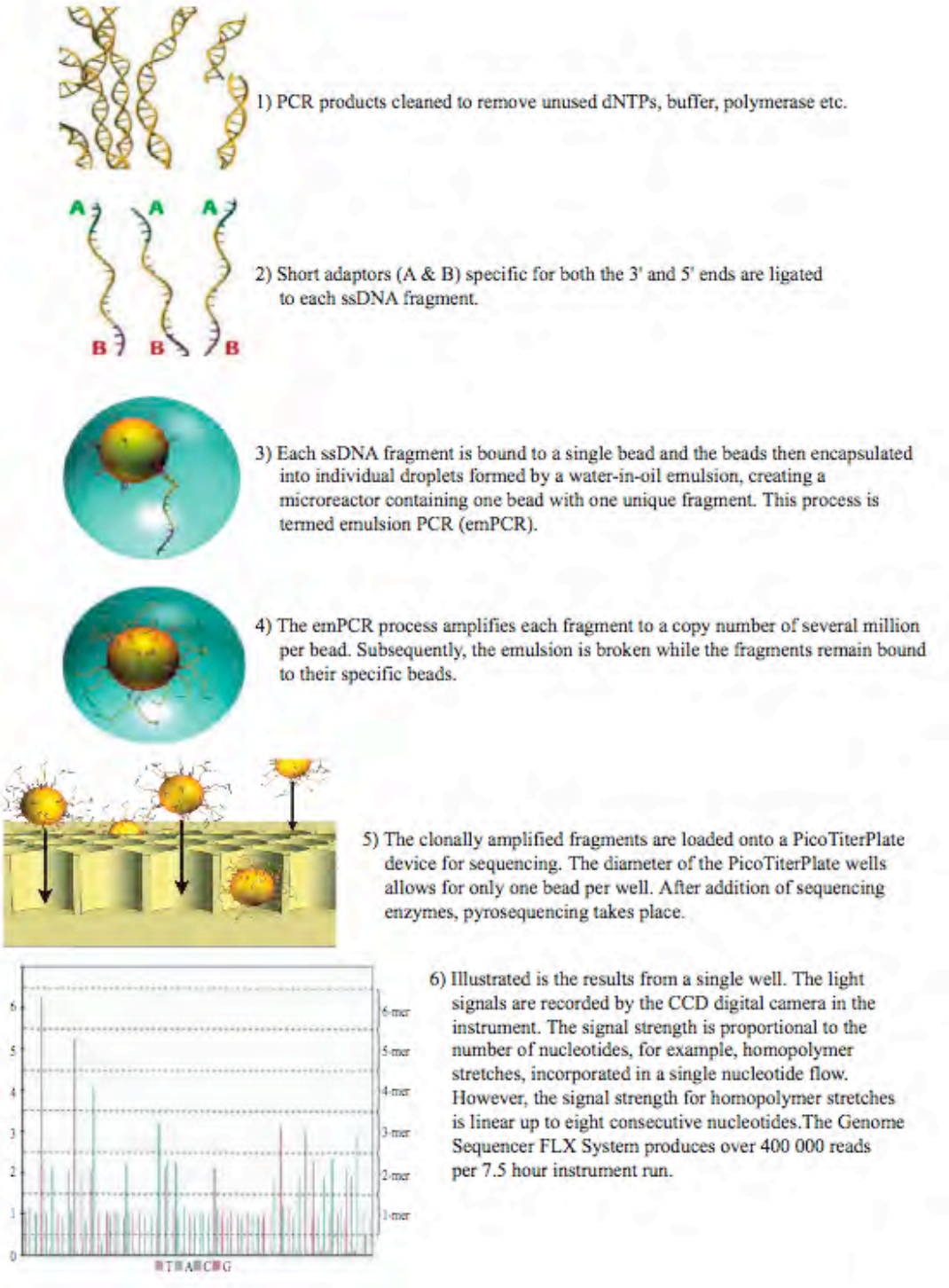


Figure 2.1. Flow diagram of Genome Sequencer FLX System

2.9 Sequence Authentication

Ancient DNA studies have been notoriously prone to contamination, with the result that strict guidelines are normally followed to ensure that each sequence is authentic (Cooper and Poinar 2000; Pääbo et al. 2004; Willerslev and Cooper 2005). To counter these problems, many sequences were repeatedly amplified, a number of samples were extracted twice in the HWABC and many were re-extracted in independent laboratories.

Chapter 3

Tracking Down the Last Mammoths:

Proceed with Caution!

At midnight in the museum hall,
The fossils gathered for a ball,
There were no drums or saxophones,
But just the clatter of their bones,
Rolling, rattling carefree circus,
Of mammoth polkas and mazurkas,
Pterodactyls and brontosauruses
Sang ghostly prehistoric choruses,
Amid the mastodonic wassail
I caught the eye of one small fossil,
"Cheer up sad world," he said and winked,
"It's kind of fun to be extinct."

Ogden Nash, Carnival of the Animals

3.1 Abstract

Woolly mammoths and many other large mammals are thought to have disappeared throughout northern Eurasia and North America by ~11 000 calendar years ago, except on isolated islands. However, macrofossils (bones and teeth) of 'last' individuals are unlikely to be discovered, particularly in dwindling populations. Because mammoth-derived DNA in permafrost, deriving from faeces, urine, skin flakes, hair and tissues (Higuchi *et al.* 1988; Taberlet *et al.* 1993; Poinar *et al.* 1998; Hofreiter *et al.* 2000; Valiere and Taberlet 2000; Gilbert *et al.* 2004; Gilbert *et al.* 2007) is likely to be far more ubiquitous in the environment, the chance of refining the last appearance date is greatly enhanced. However, this method relies on the ability to reliably date sediments, and to ensure that these sediments have not undergone remixing, or have been contaminated with asynchronous DNA.

3.2 Introduction

Woolly mammoths (*Mammuthus primigenius*) are widely considered to have disappeared from present-day mainland Eurasia ~11 000 calibrated radiocarbon (^{14}C) years before present (calBP) (Martin 2005). It is difficult to be precise about the time of extinction using 'last' appearance dates (LADs) based on radiocarbon-dated macrofossils, because bones and teeth inevitably become rarer as populations dwindle or become geographically restricted. Under these conditions, small surviving populations may be missed in macrofossil surveys. To test this possibility, we attempted to use 'sedimentary' ancient DNA (*sed*aDNA) (Willerslev *et al.* 2003) from dated contexts as a novel means of improving persistence estimates for mainland Siberian mammoths.

3.3 Sites and Sample Information

Sediment samples were taken from a total of 11 permafrost sites (Figure 3.1) situated in the high arctic of Alaska and Russia the sites being selected on the basis of their presumed Holocene age, and stability of the permafrost. Samples from nine of these eleven sites produced DNA of mammals that are typical arctic inhabitants, and samples from two sites did not produce any arctic mammal DNA.

Of the nine sites that did produce mammalian DNA, Drainage Lake (DL) Sites 1 and 2, have received the most attention because initial results suggested the persistence of mammoth deep into the Holocene.

DL Sites 1 and 2 are located on Cape Sabler, a narrow (~4 km long × 2 km wide) peninsula on the western shore of Lake Taimyr. Previous geomorphological and ^{14}C dating studies (Moller *et al.* 1999; Hubberten *et al.* 2004) have shown that Cape Sabler remained glacier-free for at least the last 40 000 years, during which interval, silt, fine sand, and peat have accumulated contemporaneously with the formation (and destruction) of ice-wedge polygons (Moller *et al.* 1999; Hubberten *et al.* 2004). Clastic sediments are derived from aeolian and lacustrine sources, the latter during periods of elevated lake level. The study sites are located on the north-facing side of an incised gully that drains into Lake Taimyr, ~270 m (Site 1) and ~70 m (Site 2) from its western shore and ~10 m above present lake level. Both sections were cleaned back to expose perennially-frozen deposits, from which we collected samples for dating and *sed*aDNA analysis.



Figure 3.1 Locations of Holocene Permafrost sites

At DL Site 1, three samples were collected from depths of ~5.6 and 4.0 m below modern ground surface, from blue-gray and light-brown sandy silts that are weakly laminated horizontally. These deposits are interpreted as paleolake sediments; mammoth DNA was recovered from one of these samples.

At DL Site 2, four samples were collected from within 3 m of the modern ground surface; mammoth DNA was recovered from three of these samples. The lowermost sample (3.0 m depth) was collected from a layer characterised by beds of blue-gray silty sand separated by thin laminations of dark-gray, organic-rich silty sand, dipping towards the lake at 10–15°. We interpret this as a lakeshore or nearshore (shallow water) depositional environment. The overlying layer consisted of horizontally-laminated light-brown and dark-grey silty sand, grading upwards into blue-gray, massive sandy silt from which two samples (2.45 and 2.1 m depth) were collected. Brown mottles are more abundant in the lower part of the sandy silt, interbedded with increasing amounts of peat higher in this layer.

3.4 Methods

3.4.1 Dating

3.4.1.1 Optically stimulated luminescence (OSL) dating

OSL dating provides an estimate of the time since luminescent minerals, such as quartz, were last exposed to sunlight (Aitken 1998; Lian and Roberts 2006). The frozen deposits were safely sampled in daylight using a portable drill with coring attachment, and kept frozen until allowed to thaw in the laboratory under safelight conditions (Aitken 1998). The melted rinds of each core were discarded and the mineral grains embedded in the unilluminated center were retained for dating. OSL

ages (Table 3.2) were calculated from the burial dose (paleodose), measured using the OSL signal, divided by the dose rate due to ionizing radiation. Concentrations of ^{238}U, ^{235}U, ^{232}Th (and their daughter products) and ^{40}K in dried and powdered samples were estimated using beta-particle counting, inductively-coupled plasma (ICP) optical emission spectroscopy, ICP mass spectrometry, and instrumental neutron activation analysis. Concordance between the beta-counting and parental radionuclide measurements indicates that significant disequilibrium is not present in the U or Th decay chains. The measured values were converted to dose rates, making allowance for beta-dose attenuation, field water content, and mass of organic matter (Lian *et al.* 1995). Cosmic-ray and internal dose rate contributions were included in the total dose rates. The calculated OSL ages assume that the measured radionuclide concentrations and water contents have prevailed throughout the period of sample burial. Paleodoses were determined from the ultraviolet OSL emissions of 90–125 and 125–180 μm-diameter quartz grains, prepared using standard procedures (Aitken 1998) and analysed using the single-aliquot regenerative-dose protocol (Olley *et al.* 2004; Lian and Roberts 2006; Wintle and Murray 2006). Several hundred individual grains of quartz were measured for each sample, in addition to multi-grain aliquots (containing ~80 grains) of core samples 3–5. Single grains were stimulated by a green laser beam for 1 s at 125 °C; paleodoses were obtained from signals that had decayed to background within the first 0.25 s, using the final 0.1 s as background. Grains were rejected if their OSL signals were depleted by >10% following infrared excitation (Wintle and Murray 2006). Multi-grain aliquots were given an infrared bleach for 100 s at 50°C and then stimulated by blue LEDs for 200 s at 125°C (Olley *et al.* 2004); paleodoses were calculated from the first 0.8 s of signal, using the final 8 s as background. The natural, regenerative and test doses were preheated at 180°C for 10 s

before optical stimulation, which recovered the correct (given) dose for five aliquots of core sample 4 (ratio of measured/given doses: 0.999 ± 0.023). Dose-response curves were fitted to the OSL data to estimate the paleodose and its measurement uncertainty (from counting statistics, instrumental reproducibility and curve-fitting errors) for grains with relative errors on the natural test-dose signal of <30%, recycling ratios consistent with unity at 1σ (or ±10% of unity for multi-grain aliquots), and recuperation values less than, or within error of, 0.5 Gy. For all samples, the relative spread in paleodose remaining after taking measurement uncertainties into account is less than, or within error of, 20%, as reported for well-bleached quartz grains that have not been affected by post-depositional mixing (Olley *et al.* 2004; Lian and Roberts 2006). The restricted range of paleodose values (3–24 Gy) for single grains and multi-grain aliquots further indicates that bleaching was sufficient at deposition, so the weighted mean paleodose for each sample was calculated using the central age model (Olley *et al.* 2004; Lian and Roberts 2006). The single-grain ages for Site 2 are statistically indistinguishable, and give a weighted mean of 5.62 ± 0.27 ka. We place greatest confidence in the ages obtained from single grains, because the OSL characteristics of each grain have been individually evaluated against objective performance criteria, and only grains considered reliable contribute to the final age determination.

3.4.1.2 Radiocarbon dating: Bulk Organic

Peat from the cores were given an acid–base–acid treatment consisting of sequential washes with 1 M hydrochloric acid (80°C, 20 mins), 0.2 M sodium hydroxide (room temperature, 20 mins) and 1 M hydrochloric acid (80°C, 1 hr), with thorough rinsing with ultra-pure water after each wash. The samples were freeze-dried before being

combusted, graphitized and the radiocarbon content determined by accelerator mass spectrometry (Arnold *et al.* 2008).

3.3.1.3 Radiocarbon dating: Macrofossil

Macrofossils are identifiable pieces of plant matter large enough to be studied without the aid of a microscope. They include seeds, leaves or fine rootlets. These were removed by hand and radiocarbon dated using standard techniques and procedures which included acid-alkali-acid (AAA) pretreatments.

3.3.1.4 Radiocarbon dating: Low Temperature Humin

Radiocarbon dating of bulk sediments has been shown to be complicated by the presence of multiple organic carbon fractions, each with a potentially different 14C activity. It has been demonstrated that the humin fraction (the insoluble carbon that remains in the sediment after an AAA pre-treatment) when analysed using a temperature stepped-combustion technique, gives results that are close to those obtained from macrofossils (McGeehin *et al.* 2001).

3.4.2 DNA sequence authenticity

When working with ancient DNA (aDNA), it is critically important to avoid false positive results due to contamination by modern DNA. Such contamination can arise from the samples themselves (sample-based contamination), during experimental procedures (laboratory-based contamination), and/or via DNA leaching and sediment disturbance (Willerslev *et al.* 1999; Cooper and Poinar 2000; Hofreiter *et al.* 2001; Willerslev *et al.* 2003; Hebsgaard *et al.* 2005; Willerslev and Cooper 2005; Hansen *et al.* 2006; Haile *et al.* 2007). The following criteria and considerations were used for authentication of the DNA results:

Sample-based contamination: To control for contamination during sampling and handling of the frozen core samples, the head of the drilling apparatus was spiked with a solution of *Serratia marcescens* bacteria (Willerslev *et al.* 2003). Upon arrival in Oxford, the outermost ~3 cm of each frozen core was removed under controlled conditions, and the internal portion used for further analysis. Using *S. marcescens*-specific primers and 45 cycles of PCR, we found evidence of vector contamination in the outermost material, but not in the DNA extracts from the interior portion. This result confirms previous studies showing that sample-based contamination can be successfully controlled in permafrost cores by keeping them frozen until processed in the laboratory (Willerslev *et al.* 2003; Hansen *et al.* 2006).

Laboratory-based contamination: Sample processing, DNA extraction, and PCR setup were carried out in dedicated aDNA facilities at Oxford and Copenhagen Universities using established aDNA precautions (Cooper and Poinar 2000; Hofreiter *et al.* 2001; Hebsgaard *et al.* 2005; Willerslev and Cooper 2005). To check for accuracy and the possibility of contamination, mammoth mtDNA sequences were independently obtained at the Oxford and Copenhagen laboratories. This confirms mammoth presence, as it is unlikely that geographically isolated facilities should suffer from the same form of contamination.

DNA leaching and sediment disturbance: Previous studies have shown that although DNA leaching may occur in unfrozen sediments (Haile *et al.* 2007), it usually does not under permafrost conditions (Hansen *et al.* 2006). To gauge the likelihood of DNA movement in environments typical of arctic Siberia, we tested modern surface

samples (from the active thaw layer) for mammalian DNA at several localities. No *Mammuthus* or *Ovibos* DNA was recovered from these samples, although sequence information distinctive of contemporary palearctic species was detected (*e.g., Dicrostonyx groenlandicus* (northern collared lemming) and *Rangifer tarandus* (reindeer)). These results suggest that remobilized ancient DNA is not ubiquitous in the modern environment. At least in low-energy contexts, genetic material appears to become quickly incarcerated in and bound to the sediments in which it was initially deposited with unbound, or free DNA being readily degraded by bacterial exo-nucleases (Cai *et al.* 2006).

3.4.4 DNA extraction and amplification

Approximately 10 g wet weight of frozen sediment was subsampled, placed in PowerMax™ Soil PowerBead tubes (Cambio UK Ltd.), and dissolved in 24 ml lysis buffer (Bulat *et al.* 2000). The tubes were then agitated vigorously for 1 min and left to incubate overnight at 65 °C under gentle agitation. Following extraction, the DNA was purified using the PowerMax™ Soil DNA Isolation Kit protocol, before being concentrated with Amicon Ultra 30 kDa spin columns. PCR was used to amplify mammal 16S, musk ox control region, mammoth control region and mammoth cytochrome b mtDNA using the primers listed in Table 3.1. Two µl of DNA extract was subjected to 55–60 cycles of PCR (1.5 min initial denaturation at 97 °C, 45 s at 94°C, 45 s at 56–60 °C, 1.5 min at 68 °C and a final cycle of 10 min at 68 °C).

Table 3.1 PCR primers and sequences, with references, annealing temperatures, and approximate product length.

Name	Sequence	Target Group/ Species	Reference	Annealing temperature (ºC)	Product length (bp)
16Smam1	5'- CGGTTGGGGTGACCTCGGA	Mammals	(Taylor 1996)	60	91
16Smam2	5'- GCTGTTATCCCTAGGGTAACT	Mammals	(Taylor 1996)		
16Smam3	5'- TGGGGTGACCTCGGAGAAY	Mammals	This study	57	78
16Smam4	5'- TCAACGGAMCAAGTTACCCTA	Mammals	This study		
Mammoth 16Sf1	5'- AAGATACAACTAAACAGCATAATAA	Mammoth	This study	56	172
Mammoth 16Sr1	5'- TTATTATGCTGTTTAGTTGTATCTT	Mammoth	This study		
Mammoth CRF1	5'- CATGCTTATAAGCAAGTACTGT	Mammoth	This study	56	165
Mammoth CRR1	5'- TGAGAAATCTCTAGTCATCATG	Mammoth	This study		
Mammoth CytB F1	5'- GAATTATTCGACAACTACACTC	Mammoth	This study	57	71
Mammoth CytB Cb-EB	5'- GGTATTTCAAGTTTCCGAGTAT	Mammoth	(Noro *et al.* 1998)		
Mox CRF1	5'- CTATTTAAACTACTCCCTGAATT	Musk Ox	This study	56	103
Mox CRR1	5'- TCCATCATGCGTTGTTGCGT	Musk Ox	This study		
Mox CytB F1	5'- CTGTAACACACATTTGCCGG	Musk Ox	This study	56	177
Mox CytB R	5'- TTGTTATGAGTAGAAGAATTACC	Musk Ox	This study		
BisCR-16633f1	5'- GCCCCATGCATATAAGCAAG	Bison	(Shapiro *et al.* 2004)	60	248
BisCR-16810r	5'- GCCTAGCGGGTTGCTGGTTTCACGC	Bison	(Shapiro *et al.* 2004)		
BisCR F1	5'- TATCCTTGRCAACATGCATAC	Bison	This study	56	147
BisCR R1	5'- TCTGATAAAGTTCATTAAATAGCG	Bison	This study		
ATP8_20F	5'- GCCACAGTTAGATACATC	Mammals	(Barnett *et al.* 2005)	50	181
ATP8_164R	5'- GAGGTGAATAGATTTTCGTTC	Mammals	(Barnett *et al.* 2005)		
Rangifer CR F1	5'- AGCATGRCAGTCAATGGTAG	Rangifer	This study	56	118
Rangifer CR R1	5'- GTCTGTTAAAAATGATRGTAACC	Rangifer	This study		

PCR products were cleaned using a QIAquick PCR Purification Kit (QIAGENE). Amplification products from the 16S primers were cloned and their sequences aligned using ClustalW in BioEdit (Hall 1999).

3.4.5 Taxonomic identifications

A Bayesian method was used to assign aDNA sequences to taxa (Table 3.2). For each aDNA sequence a set of closely related homologues was compiled using BLAST searches against GenBank and annotation from NCBI's taxonomy browser. Initially, only the best matching representative of each species was included in the set. Subsequent to the acceptance of three homologues only BLAST hits with a bit-score larger than half the value of the best hit were accepted. In accordance with these rules up to 30 homologues were included in the set. If the bit-score cut-off was not reached at this point up to 20 additional homologues were added, now only accepting hits that provided further taxonomic diversity at levels genus or higher. If the set thus compiled contained less than 50 sequences previously rejected species representatives with bit-scores above the cut-off was added to the alignment. Analysis was only continued if the set included at least one BLAST hit with an E-value below 0.1. The homologues were aligned using ClustalW and a version of MrBayes allowing topological constraints on trees was applied to the alignment using a GTR+G model. After burn in k = 9000 trees were sampled from a chain of 1 000 000 updates. The probability that an aDNA sequence (Q) belongs to taxon F can be assessed as follows:

$$\Pr(Q \in F \mid X) = \int_{G \in \Omega} I(Q, F \text{ monophyletic in } G)\, p(G \mid X)\, dG$$
$$\approx \frac{1}{k} \sum_{i=1}^{k} I(Q, F \text{ monophyletic in } G_i)$$

where W is the set of all possible trees (G) and Gi is the ith tree sampled from the Markov chain, X is all of the sequence data, and $I(Q,F$ monophyletic in $G)$ is an indicator returning 1 if Q and F are monophyletic in tree G and 0 otherwise.

3.5 Results

The samples are dominated by DNA sequences of rodents and large ungulates (Table 3.2), due perhaps to their characteristically large standing biomasses and the quantity of DNA they disperse as urine, feces, shed skin and hair. No carnivores were identified in the sediments, despite rare occurrences of wolf and arctic fox in the local paleontological record (MacPhee *et al.* 2002).

	OSL age (ka)			14C age (uncalibrated yr BP)			calibrated 14C age (yr BP)*	BLAST Species	Depth below gnd surf (cm)
Stevens Village (SV)									
DNA32				6950	±	80	7944-7626	Microtus[d]	150
DNA32				6990	±	80	7959-7676		
DNA32				7110	±	100	7728-8161		
DNA32				6950	±	70	7934-7668		
DNA32				6800	±	70	7790-7515		
DNA31									
DNA30	9.75	±	2.12					Loxodonta/Mammuthus[b, B, c], Microtus[d], Bison[d]	290
DNA29				9210	±	25	10488-10261	Microtus[c]	400
DNA28 + 60cm	11.35	±	2.32						
DNA28									550
DNA27				8320	±	200	9694-8648	Microtus[d], Ovibos[c]	600
DNA26 + 30cm	12.00	±	1.77						
DNA26									710
DNA23				9752	±	85	11327-10782	Microtus[d]	800
DNA23				9660	±	110	11247-10698		
DNA23				9450	±	120	11166-10309		
DNA22				9831	±	130	11756-10782		880
DNA25				9813	±	91	11609-10805		950
DNA21									
DNA20									105
DNA19									115
DNA18									128

Purgatory (PU)									
Waypoint 17 Sample 16									
Waypoint 17 Sample 15									
Waypoint 17 Sample 14									100
Waypoint 17 Sample 13									200
Waypoint 17 Sample 12									300
Waypoint 17 Sample 11 +3m	8.72	±	0.91						
Waypoint 17 Sample 11 +20cm				10270	±	25	12144-11841		
Waypoint 17 Sample 11	9.41	±	1.00					Mammuthus[d]	720
Waypoint 17 Sample 10									750
Waypoint 17 Sample 9									750
Waypoint 17 Sample 8				10710	±	60	12851-12653		760
Waypoint 17 Sample 7									790
Waypoint 17 Sample 6									800
Waypoint 17 Sample 5 +63cm	8.37	±	0.85						820
Waypoint 17 Sample 4									840
Waypoint 17 Sample 3	11.29	±	1.03	11420	±	60	13401-13173		900
Waypoint 17 Sample 2									925
Waypoint 17 Sample 1									950

	OSL age (ka)			14C age (uncalibrated yr BP)			calibrated 14C age (yr BP)*	BLAST Species	Depth below gnd surf (cm)
Khatanga 1 (Kh1)									
Khatanga 1A Site 1	3.32	±	0.33					Microtus[d]	
Khatanga 3A Site 1									
Khatanga 4A Site 1								Ochotona[d]	
Khatanga 5A Site 1								Ochotona[d]	
Khatanga 7A Site 1	7.00	±	0.51					Ochotona[d], Alces[e]	
Khatanga 2 (Kh2)									
Khatanga 1A Site 2	1.29	±	0.19					Ovibos[d], Rangifer[e]	
Khatanga 2A Site 2								Microtus[d]	
Khatanga 3A Site 2									
Khatanga 4A Site 2									
Khatanga 5A Site 2	1.22	±	0.28						
Bol'shaya Balakhnaya 05 (BB05) 73o39.176N. 102o07.289E									
Bol'shaya Balakhnaya 1B Site 1	1.95	±	0.15					Mammuthus[a, A, C], Microtus[c], Ochotona[c], Dicrostonyx[C], Ovibos[C], Lemmus[C]	
Bol'shaya Balakhnaya 2B Site 1								Dicrostonyx[c], Microtus[d], Ovibos?[d]	
Bol'shaya Balakhnaya 3B Site ?									
Bol'shaya Balakhnaya 4B Site 1									
Bol'shaya Balakhnaya 5B Site 1				32130	±	350	38500-35000**	Lepus[d], Alces[e]	
Bol'shaya Balakhnaya 6B Site 1	44.98	±	5.63					Loxodonta/Mammuthus[b, d], Microtus[d], Equus[f]	
Bol'shaya Balakhnaya 7B Site 1									
Bol'shaya Balakhnaya 8B Site 1								Mammuthus[a, c], Lepus[d], Bison[e, j], Ovibos[h], Rangifer,[i] Microtus[f]	
Bol'shaya Balakhnaya 9B Site 1								Mammuthus[c]	
Bol'shaya Balakhnaya 10B Site 1								Bison[d], Ovibos[h]	
Bol'shaya Balakhnaya 11B Site 1	49.97	±	10.55					Ovibos[g]	
Bol'shaya Balakhnaya 12B Site 1									
Bol'shaya Balakhnaya 13B Site 1	56.82	±	6.78					Bison[d, e], Muntiacus[d]	
Bol'shaya Balakhnaya 06 (BB06) 73o39.176N. 102o07.289E 34m amsl									
Bol'shaya Balakhnaya A								Dicrostonyx[d], Rangifer[d].	Surface 33
Bol'shaya Balakhnaya B									Non frozen sub surface 77
Bol'shaya Balakhnaya E	1.96	±	0.20						165
Bol'shaya Balakhnaya F	2.28	±	0.16						240
Bol'shaya Balakhnaya G	34.81	±	2.28					Mammuthus/Loxodonta[b]	275
Bol'shaya Balakhnaya H								Mammuthus[b, d], Ovibos[h]	300
Bol'shaya Balakhnaya I								Mammuthus/Elephas/Loxodonta[b]	335

	OSL age (ka)			14C age (uncalibrated yr BP)			calibrated 14C age (yr BP)*	BLAST Species	Depth below gnd surf (cm)
Buor Khaya 1 (BKh1) 71o54.341N; 132o 47.118E 14m amsl									
Buor Khaya site 01 A1	3.91	±	0.30						Surface
Buor Khaya site 01 F2	3.81	±	0.46						117
Buor Khaya site 01 G3	no osl sample recovered			8097	±	40	9088 - 9009	Mammuthus[a,b], Equus[d], Dicrostonyx	?
				50600****	±	1000	*too old to calibrate*		
Buor Khaya site 01 H3	7.08	±	0.96	4697	±	32	5532 - 5350		152
Buor Khaya site 01 I3	6.36	±	0.64						227
Drainage Lake Site 1 (DL1) 74o 32.172N; 100o 30.067E									
Drainage Lake Site 1 A								Lepus[c]	350
Drainage Lake Site 1 B	6.6	±	1.20						400
Drainage Lake Site 1 C	7.5	±	0.80	13610	±	55	16650 - 15800	Mammuthus[c], Dicrostonyx[c,d]	560
Drainage Lake Site 2 (DL2) 74o31.829N; 100o32.020E 15m amsl									
Drainage Lake Site 2 A	5.26	±	0.49	4465	±	31	5288 - 4972	Mammuthus[c]	180
				6675***	±	35	7610 - 7480		
Drainage Lake Site 2 B	5.48	±	0.44	117105***	±	50	13707-13423	Mammuthus[b], Rangifer[d], Dicrostonyx, Ovibos[h], Lemmus[d]	210
Drainage Lake Site 2 C	5.33	±	0.39	155255***	±	60	18930-18710		245
Drainage Lake Site 2 D	5.34	±	0.42	4611	±	34	5470 - 5080	Mammuthus[c, d], Dicrostonyx[d], Lemmus[d], Ovibos[h]	
				4535	±	35	5320 - 5050		
				136105***	±	55	16610-15833		300

a=MammothCytBF1, Cb-ED; b=MammothCR F1, R1; c=16S1/2; d=16S3/4; e=BisCR-16633f, BisCR-16810r; f=ATP8_20F, ATP8_164_R; g=Mox CytB F1, R1; h=Mox CR F1, R1, i=Rangifer CR F1, R1; j=Bison 593-CR F1, R1
Caps Letters represent gene amplified in Copenhagen

* 14C calibration using calib v5.01. the age range shown is the 2 sigma calibrated range, (probability = 100%, i.e. combining two or more potential calibration ranges, where they exist)
**beyond range of INTCAL06 - calibrated using Fig4 of Hughen et al 2006 (1 sigma range)
Radiocarbon dates obtained from macrofossils, unless otherwise indicated by ***, which signifies Low Temperature Bulk Humin dating
****Radiocarbon dating on twig
Taimyr Lake Holocene and Buor Khaya site 03 samples were also extracted, but yielded no arctic mammal species.

3.6 Statistical modelling of Taimyr macrofossil dates

All published radiocarbon dates (n = 65) on macrofossil remains of mammoth that were less that 20 000 calBP (calibrated using CalPal2007; http://www.calpal.de/; max = 29 980) were collected, with a youngest age (y) of 11 020 calBP. This macrofossil sample gives a maximum likelihood estimate of the lower confidence bound of extinction (Solow *et al.* 2006) of 10 217 calBP, assuming a constant population size prior to extinction. However, if we assume that: (i) post-Pleistocene mammoth population density was suppressed for an extended period of time, before eventual extinction at min = 5600 BP, (ii) the rate of macrofossil deposition is directly proportional to density, and (iii) due to permafrost conditions there was no change in the rate of preservation over time, then we can estimate the reduction in density required to meet the condition of sampling no macrofossils between y and min at a probability α:

$$D_r = \left(\frac{max - y}{max + (\alpha - 1)y - \alpha \cdot min)} \right)^n$$

This gives a reduction required to observe a 10 230 cal BP date of 10.1 fold reduction for 95% confidence bounds or a 43.8 fold reduction for the 50% confidence bounds.

3.7 Discussion

The most recent bone-derived dates for mainland mammoth in Asia are 10 882-11 165 calBP; 11 196-11 233 calBP; and 11 234-11 317 calBP (Sulerzhitsky 1997; Sulerzhitsky and Romanenko 1997) and are from the Taimyr Peninsula (north central Siberia). The youngest Asian island dates come from Wrangel Island; 4014-4144 calBP, 4308-4412 calBP, 4438-4555 calBP and 4457-4558 calBP (Vartanyan *et al.*

1993; Long *et al.* 1994; Vartanyan *et al.* 1995; Kuzmin *et al.* 2000). In North America, the most recent mainland date is 13 220-13 565 calBP and most recent island date is 8628-8933 calBP from St Paul Island (Guthrie 2004).

Table 3.3 shows the mammal DNA results in relation to the last appearance dates of the relative fauna. The youngest macrofossil-based 14C ages for mammoths in mainland Eurasia, of 11 150 yr BP, occur in the Taimyr peninsula of north-central Siberia (MacPhee *et al.* 2002), suggesting that it acted as a temporary refugium during the mammoths' decline toward extinction. Therefore the existence of mammoth DNA from Bol'shaya Balakhnaya 05 (45 000 years), Bol'shaya Balakhnaya 06 (35 000 years), Drainage Lake Site 1 (16 000 years), and Drainage Lake Site 2 (>13 000 years) from this area is not unexpected.

Table 3.3 Showing mammal DNA results in relation to last appearance dates

	***Mammuthus* (Ex)**	***Equus* (LEx)**	***Ovibos* (LEx)**	***Bison* (LEx)**	***Alces* (LEx)**	***Rangifer* (L)**	***Lemmus* (L)**	***Dicrostonyx* (L)**	***Microtus* (L)**	***Ochotona* (L)**	***Lepus* (L)**	***Muntiacus* (L)**
LAD, cal yr BP	Mainland Asia: 11000	2140	2780	10040	4700							
	Mainland America: 13300											
	Wrangel Island:											
	Pribiloff: 6500											
<2000	BB05		Kh2, BB05			Kh2, BB06	BB05	BB05, BB06	Kh2, BB05	BB05		
<3000												
<4000									Kh1			
<5000												
<6000											DL1	
<7000	BKh1	BKh1			Kh1			BKh1	SV	Kh1		
<10,000	SV, PU, DL2		SV	SV					SV			
>10,000	BB05, BB06 , DL1, DL2	BB05	BB05, BB06, DL2	BB05	BB05	BB05, BB06, DL2	FL2	DL1, DL2	BB05, SV		BB05	BB05

Colored cells show position of species LADs. Ex, extinct; LEx, locally extinct; L, living.
BB05, Bolshaya Balakhnaya 2005 cores; BB06, Bolshaya Balakhnaya 2005 cores; BKh, Buor Khaya site #1; DL1, Drainage Lake site #1; DL2, Drainage Lake site #2 , Cape Sabler, site #2; Kh1, 2, Khatanga River sites #1, #2; PU, Purgatory; SV, Stevens Village
Codes in red represent potentially significant samples.

On the other hand, the presence of mammoth DNA from Bol'shaya Balakhnaya 05 from sediments dating to 2000 yr BP is highly improbable because the 16S mammoth sequences from this site come from Clade II mammoths. The most youngest dated fossils of Clade II mammoths are 40-44 000 years old (Barnes *et al.* 2007; Gilbert *et al.* 2008). This site of Bol'shaya Balakhnaya 05 lies on the bank of a river and the river cuts through deposits that are over 45 000 years old (Table 3.2). It is therefore likely that reworking and redeposition of Pleistocene sediments, with Pleistocene mammoth DNA bound to it, has occurred, "contaminating" the Holocene layers. In spite of this reworking, Holocene OSL dates are generated due to the re-zeroing of the quartz crystals during redeposition events.

Laboratory experiments using samples which were known to contain mammoth DNA, and which were then agitated with DNA-filtered Siberian river water, did not release the DNA into the water and mammoth DNA remained bound to the sediments, which was subsequently extracted. The unfiltered river water, which flows beside a section containing DNA-bearing Pleistocene deposits (see Chapter 4), did not yield any arctic mammal species. This finding suggests that aDNA contamination of modern sediment samples by old DNA is caused by bulk movement of sediments, rather than water movement.

Although the stratigraphic fidelity of mammoth DNA results from this site is unreliable, the occurrence of Clade II sequences still provides useful information; individuals from this clade had previously not been found west of the Lena River (Gilbert *et al.* 2008).

By contrast to the results of Bol'shaya Balakhnaya 05 sediments from the site of Drainage Lake Site 2 are Clade I mammoths (Barnes *et al.* 2007; Gilbert *et al.* 2008). Samples yielded internally consistent OSL ages from individual grains of different sizes and from aliquots composed of ~80 grains, both within and between DL Sites 1 and 2. Among these were several sequences distinctive of mammoth, from deposits with a weighted mean age of 5600 ± 300 years BP. The OSL dates were corroborated by three macro-fossil based dates that gave an age of approximately 5200 years BP. The site itself had been geologically interpreted as being a paleolake with the sediment beds of different content, such as peat and sand, being discrete and laminated and there being little or no evidence of heave or cryoturbation. Additionally it was discovered during the radiocarbon pretreatment procedure that very little humic material was removed by the base wash, consistent with the inference that there was negligible movement of soluble organics at this site.

However, a re-evaluation of the site (instigated by reviewers' comments), led to its reinterpretation as thermokarst depression, which is formed when ice-rich permafrost is disturbed, water is released from the excess ice, sediments collapse, heat is trapped in the water, providing a positive feedback cycle, and the depression expands by melting the basin-margin sediments, releasing water and sediment to the basin. Usually this results in retrogression (ice-rich erosion) along the margins, eroding surrounding sediments. This reinterpretation led to a decision to conduct bulk sediment dating using the Low Temperature Humin method, which returned Pleistocene dates for all the DL samples, with the exception of DL 2A. As humin is an insoluble compound, these results demonstrate that reworking of the sediments has taken place (rather than water-borne old carbon contamination), at least for lower strata. Whilst there is no evidence that DNA binds to humin (although it is known to

attach to humic acid (Crecchio and Stotzky 1998)), the stratigraphic fidelity of the site is questionable, and even though three dating techniques suggest that DL 2A is Holocene, the evidence of the underlying sediments undermines any certainty.

The site of Buor Khaya is similarly unreliable. Like DL, OSL dates place the mammoth DNA-bearing stratum well within the Holocene. Unfortunately when a radiocarbon date was obtained from an associated twig, it came back as being 50 000 years old, although plant roots from the sample gave an age of 9009-9088 calBP. Buor Khaya lies on the edge of a lake, and after being eroded from a Pleistocene deposit, the twig may have floated on the lake and become incorporated within the younger sediments. Whilst it is unlikely for ancient DNA to also be transported in this way, this and the apparent age inversion of the sediments give cause for doubt.

There remains three sites that were studied, that contain DNA-bearing sediments that have been dated as being more recent than the bone-derived LADs. Two from North America, Stevens Village and Purgatory produced mammoth DNA from a strata OSL dated to 9750±2120 BP and to 9410±1000 BP respectively. Furthermore, there are no Pleistocene deposits near or underlying the Stevens Village site (Duane Froese *pers. com.*). The Purgatory site has produced spruce macrofossils and pollen at ca. 10300-10700 14C yrs BP from a paleosol, which represents the oldest spruce macrofossils in Alaska (associated with the Holocene) and provides excellent evidence for a spruce refugium (Zazula *et al.* 2006). The third site is Khatanga, which gave musk ox DNA from a sediment dated to 1290±190 BP, a few hundred years more recent than the bone-derived LAD for this species (Sulerzhitsky and Romanenko 1997). However,

until the OSL dates are confirmed by radiocarbon dates, this finding will remain contestable.

Nevertheless, despite the problems of permafrost stability revealed by this study, it has been demonstrated that retrieval and identification of genetic material from perennially-frozen, well-dated depositional contexts can be used to establish species identifications and LADs, independent of macrofossil recovery. Direct dating of macrofossils has shown that mammoths survived on several Beringian islands until ~6500–4000 calBP (Veltre in press), many millennia after their assumed demise on the mainland. Their longer persistence on islands has generally been attributed to isolation, but this interpretation needs reappraisal: the mammoth-associated ages from Stevens Village, Purgatory and possibly also Buor Khaya, imply that arctic refugia existed in places. This observation would be compatible with current macrofossil evidence from Taimyr and Alaska if mammoth abundance declined before extinction.

Although it is likely that still younger mammoth skeletal material will be found, the systematic search for sedaDNA provides a powerful new approach for obtaining special and temporal information for recently extant mammoth populations. Whether mammoth extinction was ultimately forced by progressive changes in vegetation, increased anthropogenic pressures, or a combination of these and other factors, the process was evidently protracted in at least some continental areas, contrary to the expectations of the ‘blitzkrieg’ hypothesis for megafaunal extinction (Martin 1984).

Chapter 4

DNA-Based Biodiversity Assessment: Plant DNA from Arctic Sediments

Where there's muck there's brass.

Traditional Northern English Saying

4.1 Abstract

Assessing the biodiversity of an environment has traditionally been achieved through the identification and mapping of whole organisms. Whilst DNA-based biodiversity assessments have previously been carried out in aquatic environments (von der Heyden and Cavalier-Smith 2005; Sogin et al. 2006), here we present a novel approach to assess biodiversity of plant communities in terrestrial Arctic environments. Moreover, we compare these results to whole organism and pollen surveys in order to establish the potential of this technique to the assessment of modern plant diversity and distribution. In addition we show that the relationships between modern plant, pollen, insect and DNA distribution allows this approach to be used to investigate *past* plant diversity and distribution. This has enormous implications to fields such as palynology and palaeoecology as well as climate change investigations.

4.2 Introduction

Total plant DNA present within the rhizosphere originates from leaf litter, pollen grains, roots and rootlets. Upon cell death, the DNA is released into the environment where it is either metabolised by microorganism exonucleases, transformed through integration and uptake into competent soil bacteria or it endures for some time by binding to humic acid and soil minerals (Levy-Booth et al. 2007). The majority of such extracellular plant sedimentary DNA is thought to derive from sloughed off root cap cells (Hawes 1990; de Vries et al. 2003) and thus will be a reflection of the local plant diversity. This contrasts with pollen records which are a function of pollen dispersal as well as proximity to the source. Although DNA is present in pollen grains (Parducci and Petit 2004) and occasionally in large amounts (Turbin 1965) this

contribution is seasonal and as a proportion of total extracellular plant soil DNA considered to be negligible. In high latitudes, this proportion is presumably even smaller as there is relatively less pollen production than in lower latitudes, and root cap derived soil DNA is more plentiful as bacterial DNases, the primary mechanism for extracellular DNA degradation (Blum *et al.* 1997), are less reactive at the lower soil temperatures.

4.3 Sites and Sample Information

Situated in the Norwegian tundra, the University of Oslo's field site in Varanger Peninsula, Norway (70° 19' N, 30° 01' E), consists of arctic-alpine heath and grassland. A total of four study areas, two in Vestre Jakobselv and two in Komagdalen were each divided into two 15×15 m plots which were in turn subdivided into 13 subplots measuring 0.5×0.5 m within which biomass per vascular plant species was recorded. From inside each 15×15 m plot soil samples were taken three times from three 0.5×0.5 m subplots (Figure 4.1).

In order to test the applicability of this method for environmental reconstruction, six samples from the palaeontological site of Ice Bluff Cliff, Mayn River, Siberia (64° 17' N; 171° 15' E) (Figure 4.2) were subjected to cpDNA anaysis.

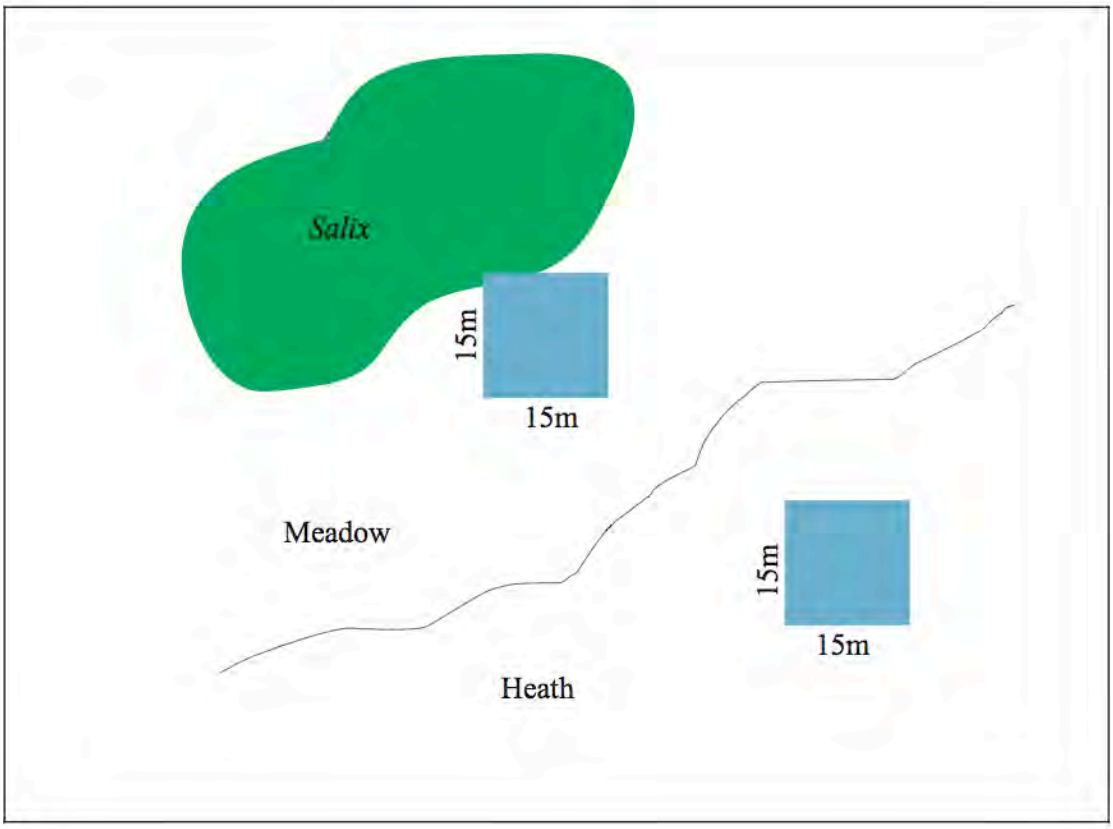


Figure 4.1 Showing a schematic design of one (of four) pairs of 15×15 m plots each containing 13 0.5×0.5 m botanically surveyed subplots, and three 0.5×0.5 m plots from where three soil samples were taken.

4.4 Methods

4.4.1 DNA Extraction and Amplification

Total DNA was extracted from approximately 6 g of soil using a combined Sergey Bulat protocol (Bulat et al. 2000) and Cambio PowerMax Soil kit protocol, which employs a silica clean-up method. Chloroplast *trn*L DNA was selectively amplified using *trn*L*g* and *trn*L*h* primers (Taberlet et al. 2007) which had been 5' labelled with a sequence of six unique bases, terminating in a cytosine base, and ligated to 454 sequencing primers.

The trnLg/h area of the chloroplast genome is flanked by highly conserved regions which allow the use of "universal" primers, and yet amplifies a region that is sufficiently variable for high resolution at the genus and species level for most taxa (Figure 4.3). In addition, the region is short, typically less than 100 bp, which is ideal for use on degraded aDNA.

When the amplified products were subsequently sequenced using a Roche Applied Science 454 Genome Sequencer, the unique tags enabled the origin of the samples to be determined (Binladen et al. 2007). A mean of 1600 sequences were generated for each modern soil sample, and a mean of 3953 sequences generated for each Pleistocene sample. These sequences were then compared to a *trn*L database of 856 species of arctic plants (289 genera in 84 families) (Taberlet *et al. in prep.*).

The residue sediment from which the DNA had been extracted was then examined for pollen grains *i.e. the exact same sample was processed for DNA and pollen*.

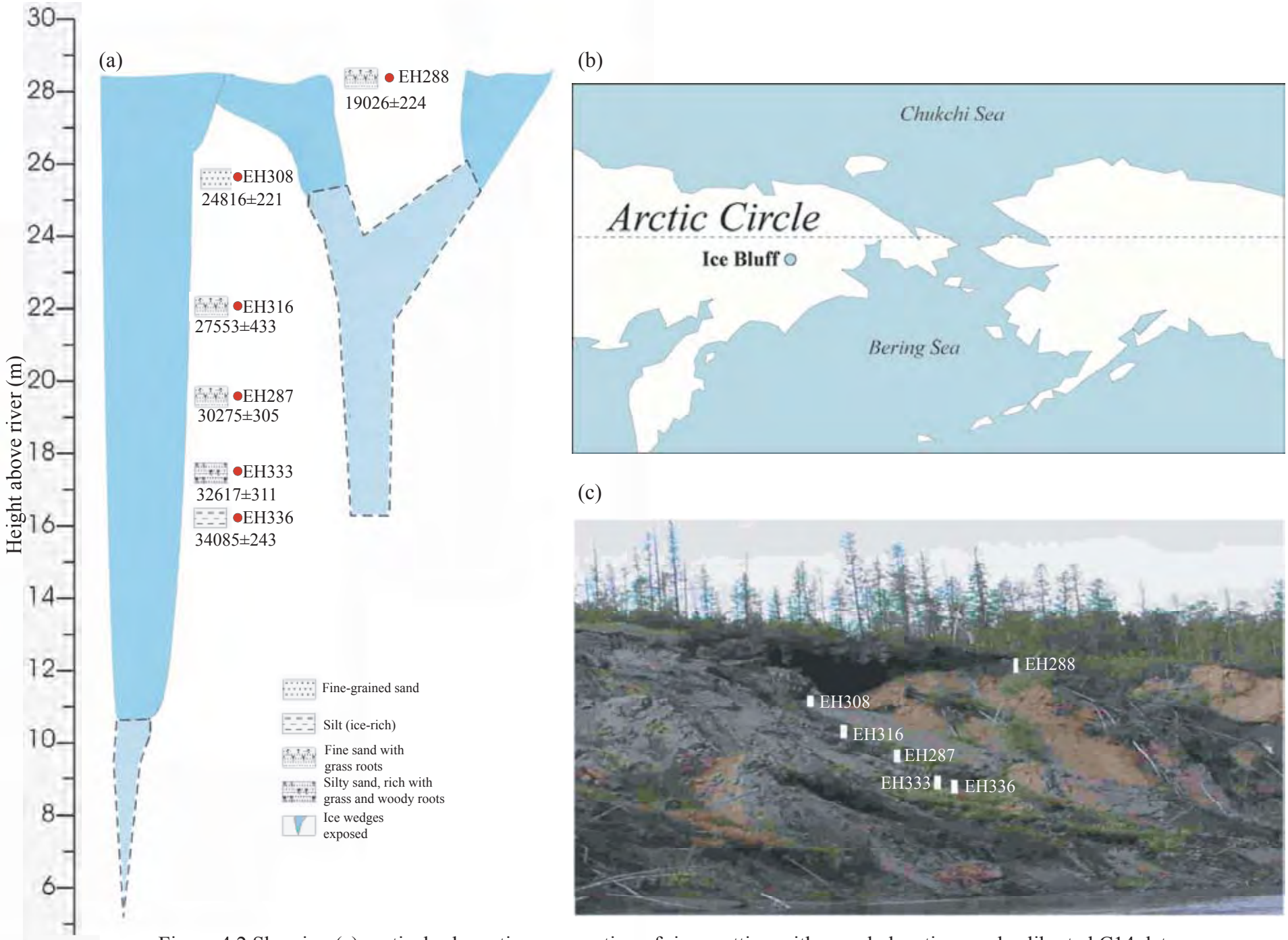


Figure 4.2 Showing (a) vertical schematic cross section of river cutting with sample locations and calibrated C14 dates, (b) location of Ice Bluff Cliff in Siberia, (c) photograph of Ice Bluff Cliff with sample locations

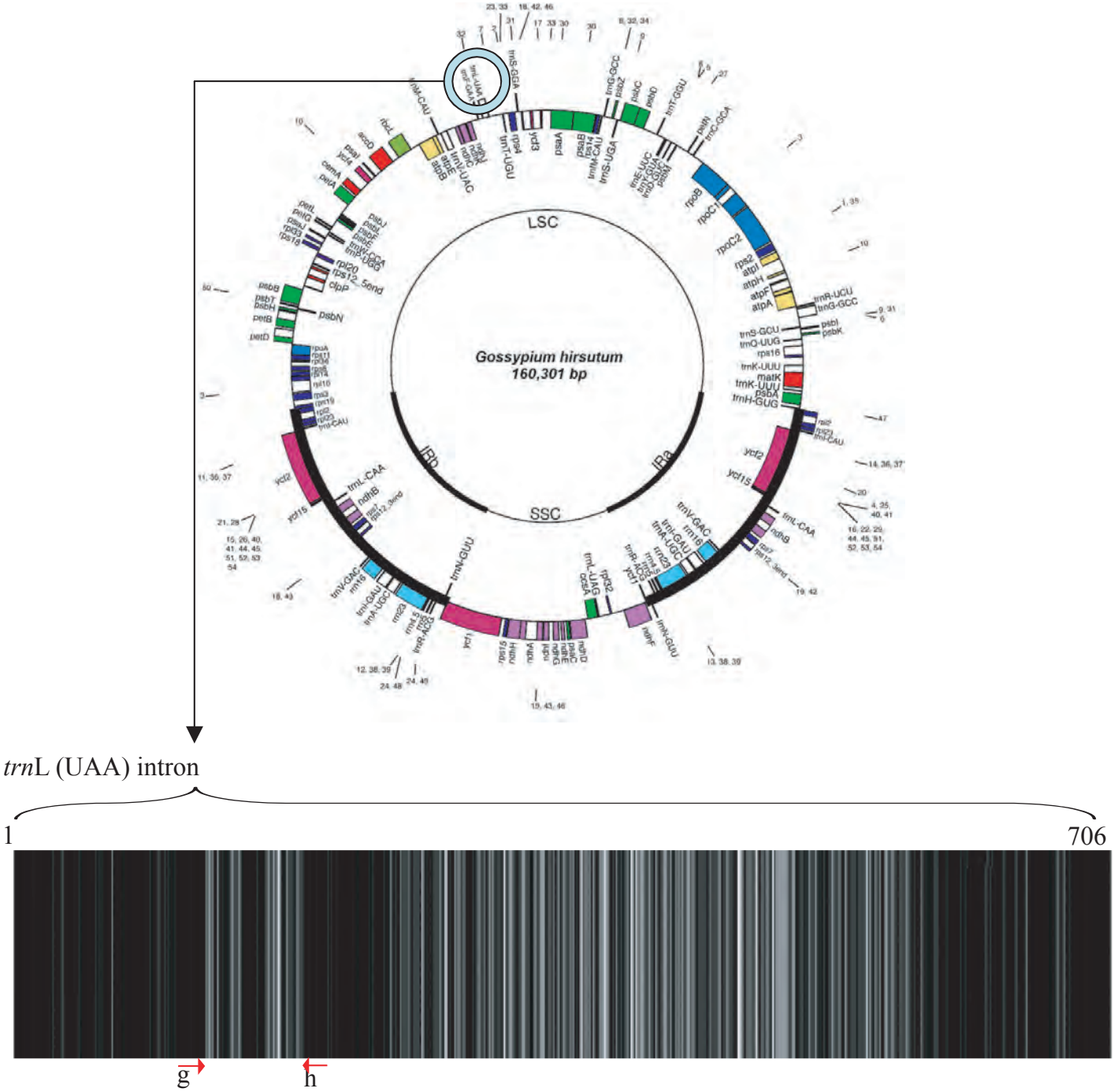


Figure 4.3 Position of the primers *g* and *h* on the chloroplast *trn*L (UAA) gene.

4.4.2 Pollen Analysis

Pollen analysis was carried out from moss pollsters, as well as from the DNA-extracted sediments on the following samples: HEI A1.2, HEI A3.1, HEI B1.2, HEI D2.1, ENG A1.2 ENG B 1.3, ENG D1.1 (Table 4.1).

To aid pollen identification a species list of known pollen types was developed by incorporating those of Prentice (Prentice 1980) and categorising an arctic species list into pollen types according to Moore et al. (Moore et al. 1991).

Slides were counted at 400× magnification and a minimum of 500 grains were counted where possible.

Table 4.1 Pollen count for Varanger Sediment Samples and Moss Pollsters

	Soil	Moss	Soil	Moss	Moss	Moss	Soil	Moss	Soil	Moss	Soil	Moss	Soil	Moss	Moss	Soil
	HEI A 1.2	**HEI A 1**	**HEI A 3.1**	**HEI A 3**	**HEI B 2**	**HEI C 3**	**HEI D 2.1**	**HEI D 2**	**ENG A 1.2**	**ENG A 1**	**ENG B 1.3**	**ENG B 2**	**ENG D 1.1**	**ENG D 1**	**HEI C 1**	**HEI B 1.2**
***Alnus* undiff.**	9			1	12	7					2		2			1
***Aster* type**	27					1		1	13	2	1	1	87	7	1	6
Betula	97	53	244	172	202	162	178	110	113	40	128	28	37	51	66	262
Campanula																1
Caryophyllaceae (c.f. silene schleranthus)				1	1		1		5	2	2		1			
Corylus									1		1					
Cyperaceae	53		2		4	5	2		12		2	1	6	40		14
***Epilobium* type**									1							
Ericales	19	32	190	191	64	107	126	228	25	6	4	7	12	13	76	110
Lactuceae	10	3			1		2	1	10	1	3	1	8	2		
Lemna																2
Lycapodium undiff			2	7	14	3	93	2	15	1			9	2		
Monolete undiff.	72		3	5	4	10	2	3	21	1	1			1	13	8
***Myriophyllum* (4 pore *Betula*)**			1	4	3		2		1		2					1
Myrica														1		
Picea			2		**2**	3		4	9	1	3		10	1.5	1	
Pinus	62	37	39	123	124	165	91	133	142	46	51	38.5	127	147	62	36
Poaceae	162	27	80	28	45	77	46	15	52	441	235	403	210	189	12	60
Polygonum aviculare		4											3			
***Ranunculus* und**	12	10	3	4	16	7	4	1	41	28	40	11	35	59	3	21
Rumex				1	1	1				2				6		
Salix	17		4		13	1	10		28	8	7	2	3			7
Selaginella selaginoides	3	1							12		41	62				
***Trilete* undiff.**	8	1	2		1	2	5	2	5	1		4		2	4	5
Trifolia (c.f.)														7		
unknown/ unidentifiable	3	2	1						3		1	1		2		5
Urtica (c.f.)						2			4			3		2		
TOTALS	**554**	**170**	**573**	**537**	**507**	**553**	**562**	**500**	**513**	**580**	**524**	**562.5**	**550**	**532.5**	**238**	**539**

4.4.3 Dating

Dating of soil samples was carried out on plant rootlets by radiocarbon analysis at the University of Oxford Radiocarbon Laboratory.

4.5 Results and Discussion

4.5.1 Botanically-Identified Plant Species Diversity

Botanically-identified plant species diversity (plant data) and abundance was measured in all of the plots (Table 4.2) and at three different sites within each plot three samples of topsoil were collected (see Appendix 1 for species codes). The estimation method applied to calculate abundance of each species is point intercept frequency which is found to correlate highly with biomass (Bråthen and Hagberg 2004).

4.5.2 DNA-Identified Plant Species Diversity

DNA-identified plant species diversity (DNA data) results from modern and ancient permafrost are presented in Appendices 2 and 3. The DNA data from the topsoil is likely to reflect modern, or near-modern plant communities, as reworking of old material is considered unlikely in this low-energy environment (Fontaine et al. 2007).

Table 4.2 Botanically Identified Plant Species surveyed at Varanger and Point Intercept Frequency (Abundance). ENG= Grassland, HEI=Heath.

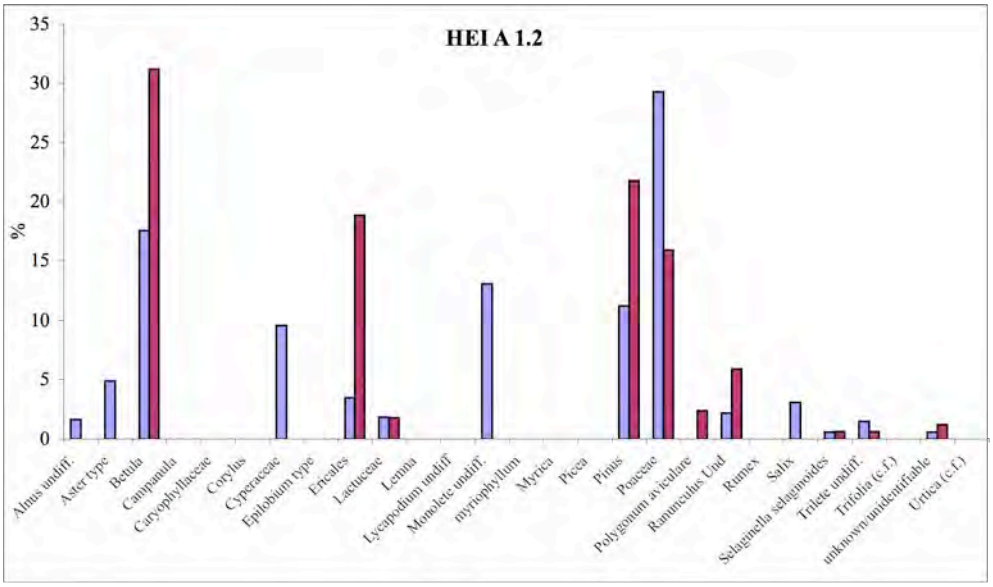
ENG A		HEI A		ENG B		HEI B		HEI C		ENG C		HEI D		ENG D	
Species	**Abundance**	**Species**	**Abundance**	**Species**	**Abundance**	**Species**	**Abundance**	**Species**	**Abundance**	**Species**	**Abundance**	**Species**	**Abundance**	**Species**	**Abundance**
Des fle	489.0	*Vac myr*	172.0	*Des fle*	316.0	*Des fle*	190.1	*Vac myr*	106.1	*Car aqu coll*	143.0	*Vac myr*	165.1	*Des fle*	217.1
Des ces	153.0	*Emp nig her*	136.1	*Ant odo alp*	215.0	*Bet nan*	139.1	*Bet nan*	103.2	*Car bru*	54.0	*Bet nan*	105.1	*Vio bif*	77.2
Cal pur	121.0	*Des fle*	135.0	*Des ces*	173.0	*Vac myr*	127.0	*Emp nig her*	100.0	*Ant odo alp*	48.0	*Des fle*	81.1	*Sal gla coll*	50.0
Vio bif	82.0	*Bet nan*	108.2	*Poa pra alp*	78.0	*Emp nig her*	83.1	*Cor suec*	56.0	*Rum ace coll*	33.1	*Emp nig her*	65.1	*Ant odo alp*	49.1
Ant odo alp	80.0	*Vac vit*	36.2	*Poa sp*	60.0	*Cor suec*	33.3	*Vac uli*	42.1	*Equ arv*	29.3	*Sal her*	43.5	*Car big*	43.1
Cal str	75.0	*Cor suec*	29.2	*Rum ace coll*	57.0	*Nar stri*	25.1	*Des fle*	33.1	*Car sp*	27.1	*Cor suec*	14.3	*Sal sp*	42.0
Rum ace coll	31.1	*Vac uli*	18.0	*Vio bif*	55.1	*Vac uli*	24.0	*Vac vit*	26.1	*Poa pra alp*	26.0	*Vac uli*	12.0	*Tro eur*	39.2
Poa pra alp	29.1	*Fes rub coll*	8.1	*Bis viv*	31.2	*Vac vit*	20.1	*Jun tri*	6.1	*Des fle*	24.0	*Vac vit*	11.4	*Sal her*	32.1
Sal gla coll	22.1	*And pol*	8.0	*Fes rub coll*	23.0	*Sal her*	6.4	*Car big*	5.2	*Bis viv*	17.3	*Jun tri*	10.3	*Sol vir*	24.1
Alc sp	18.3	*Sal her*	8.0	*Cal str*	21.0	*Cal pur*	4.2	*Car sp*	0.2	*Cal str*	16.2	*Ant odo alp*	3.0	*Bis viv*	22.2
Fes rub coll	14.0	*Alc sp*	7.0	*Ran acr coll*	15.3	*Car big*	3.2	*Bis viv*	0.1	*Car lac*	13.0	*Bis viv*	2.3	*Poa sp*	22.1
Vac myr	14.0	*Sal sp*	6.0	*Sal her*	15.1	*Phy cae*	2.1	*Car bru*	0.1	*Phl alp*	12.4	*Oma sup*	2.0	*Ger syl*	21.4
Equ syl	11.0	*Equ arv*	5.1	*Vio sp*	15.0	*Fes rub coll*	2.0			*Jun fil*	8.1	*Sol vir*	0.3	*Ran acr coll*	12.3
Bis viv	10.3	*Nar stri*	5.1	*Alc sp*	9.3	*Jun tri*	1.0			*Agr mer*	7.1	*Hie Alp*	0.2	*Poa pra alp*	11.1
Ran acr coll	10.3	*Ant odo alp*	4.0	*Car sp*	9.0	*Equ arv*	0.1			*Equ syl*	7.1	*Car big*	0.1	*Alc sp*	9.1
Rhi min coll	9.3	*Des ces*	4.0	*Cal pur*	8.0	*Jun fil*	0.1			*Rub cha*	6.1	*Des ces*	0.1	*Tri spi*	9.0
Phl alp	9.0	*Bis viv*	3.0	*Sal sp*	6.4	*Rub cha*	0.1			*Vah atr*	6.0	*Hie sp*	0.1	*Vac vit*	8.0
Sal phy	8.3	*Car big*	3.0	*Tar sp*	6.3					*Luz mul*	5.3	*Phy cae*	0.1	*Pyr min*	7.4
Vac uli	8.0	*Ger syl*	3.0	*Equ arv*	5.5					*Sol vir*	4.3	*Pyr sp*	0.1	*Sal lap*	7.1
Cir hel	7.0	*Luz sp*	3.0	*Equ pra*	5.2					*Tri eur*	4.1	*Sib pro*	0.1	*Tri eur*	7.0
Equ arv	7.0	*Tro eur*	2.0	*Sal phy*	5.2					*Cal pur*	4.0			*Sal phy*	6.1
Sol vir	6.0	*Ran acr coll*	1.1	*Phl alp*	5.1					*Sal phy*	3.1			*Sib pro*	6.1
Tri eur	6.0	*Agr mer*	1.0	*Car can*	3.0					*Cer alp-coll*	2.7			*Tar sp*	5.2
Vio sp	5.0	*Oma nor*	1.0	*Sal gla phy*	3.0					*Sal her*	2.1			*Fes ovi*	4.1
Bet nan	4.0	*Sol vir*	1.0	*Tar cro*	3.0					*Cer cer*	2.0			*Jun com*	3.0
Ste nem	4.0	*Tar cro*	1.0	*Cir hel*	2.0					*Poa sp*	2.0			*Rum ace coll*	2.0
Cer fon	4.0	*Tri eur*	1.0	*Hie Alp*	2.0					*Pot pal*	2.0			*Vac myr*	1.1
Agr sp	3.0	*Bar alp*	0.1	*Pyr sp*	2.0					*Sal gla coll*	2.0			*Car lac*	1.0
Cer cer	3.0	*Cir hel*	0.1	*Cer alp-coll*	1.3					*Vio sp*	1.0			*Luz mul*	1.0
Cor suec	3.0	*Equ syl*	0.1	*Rhi min coll*	1.2					*Epi hor*	0.5			*Hie Alp*	0.2
Oma nor	3.0	*Hie Alp*	0.1	*Alc alp*	1.1					*Oma nor*	0.2			*Agr mer*	0.1
Sal sp	3.0	*Jun fil*	0.1	*Sel sel*	1.1					*Vac myr*	0.2			*Cer cer*	0.1
Ver alp alp	2.1	*Luz sud*	0.1	*Epi sp*	1.0					*Epi sp*	0.1			*Eup fri*	0.1
Ang arc arc	2.0	*Rhi min coll*	0.1	*Par pal coll*	1.0					*Fes sp*	0.1			*Oma sup*	0.1
Pot pal	2.0			*Sal lan*	1.0					*Sal lap*	0.1			*Oxy dig*	0.1
Eup sp	1.1			*Epi hor*	0.3									*Phl alp*	0.1
Poa sp	1.0			*Tri eur*	0.3									*Pyr sp*	0.1
Tar cro	1.0			*Vac uli*	0.2									*Cal sp*	0.1
Car big	0.2			*Ver alp alp*	0.2										
Cer alp-coll	0.2			*Bet nan*	0.1										
Pyr min	0.1			*Emp nig her*	0.1										
Sal lap	0.1			*Luz mul*	0.1										
Tar sp	0.1			*Luz sud*	0.1										
Sag sp	0.1			*Oma sup*	0.1										
				Pot pal	0.1										
				Sib pro	0.1										
				Ste nem	0.1										

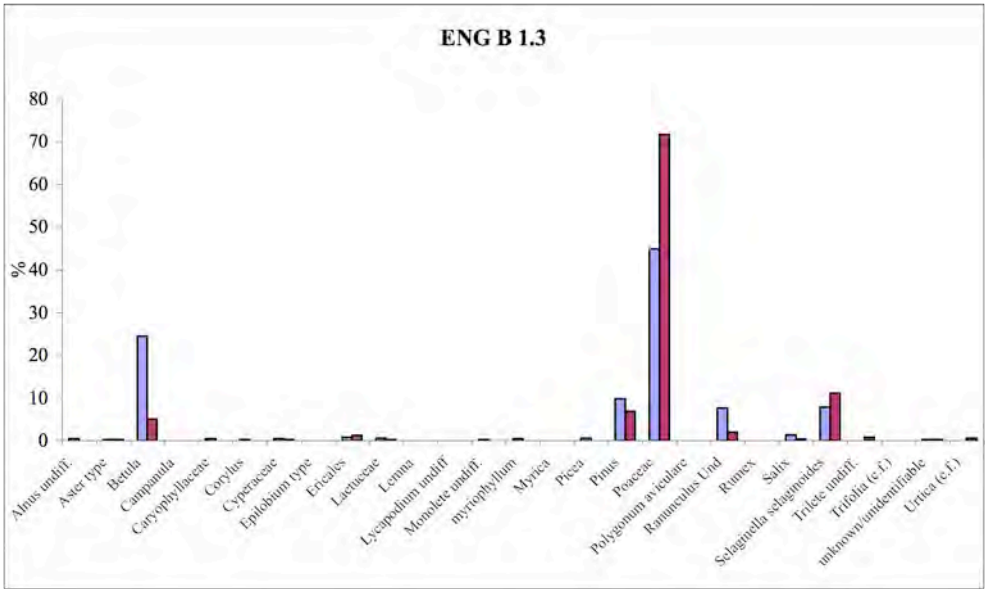
4.5.3 Pollen-Identified Plant Species Diversity

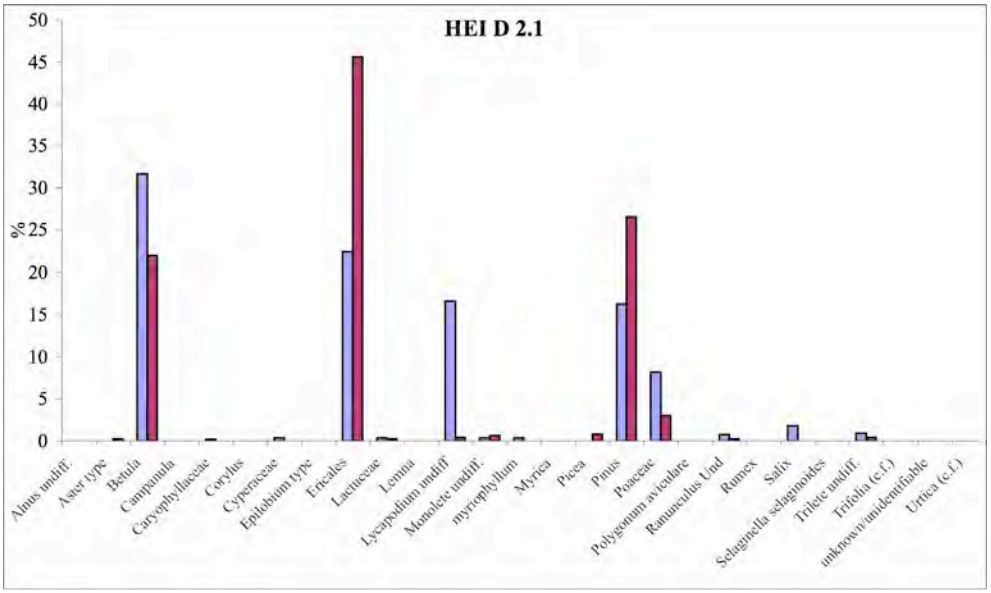
The Varanger samples generally yielded sufficient material for meaningful analysis of pollen-identified plant species diversity (pollen data). The moss pollsters were sparser than the soil samples but this was remedied by washing large samples to release the pollen. The sediments from which DNA had been extracted proved to be usable for pollen analysis and the pulverising granules appear not to have significantly damaged the pollen grains.

The relative abundance of pollen types in the moss and soil samples have been plotted as bar charts for easy comparison (Figure 4.4), excluding pine pollen which is not locally derived.

The soil samples have not been sampled stratigraphically and the age of sediments is unknown so the samples span an interval of uncertain length. It would therefore not be surprising if soil samples differed considerably in composition from the moss samples even before taphonomic processes are taken into account. In fact the data do not display this and there is generally a high degree of similarity between the two sample types. This is particularly clear when pine is discounted. The greatest differences tend to be in proportion rather than composition and dominant types tend to be found in both samples but sometimes in differing proportion. The similarity suggests that the soil samples may not be very much older than the pollen types or that the vegetation has not changed significantly over the timescale which the samples represent.


HEI A 1.2
%
35
30
25
20
15
10
5
0
Alnus undiff.
Aster type
Betula
Campanula
Caryophyllaceae
Corylus
Cyperaceae
Epilobium type
Ericales
Lactuceae
Lemna
Lycapodium undiff
Monolete undiff.
myriophyllum
Myrica
Picea
Pinus
Poaceae
Polygonum aviculare
Ranunculus Und
Rumex
Salix
Selaginella selaginoides
Trilete undiff.
Trifolia (e.f.)
unknown/unidentifiable
Urtica (e.f.)


ENG B 1.3
%
80
70
60
50
40
30
20
10
0
Alnus undiff.
Aster type
Betula
Campanula
Caryophyllaceae
Corylus
Cyperaceae
Epilobium type
Ericales
Lactuceae
Lemna
Lycapodium undiff
Monolete undiff.
myriophyllum
Myrica
Picea
Pinus
Poaceae
Polygonum aviculare
Ranunculus Und
Rumex
Salix
Selaginella selaginoides
Trilete undiff.
Trifolia (e.f.)
unknown/unidentifiable
Urtica (e.f.)


HEI D 2.1
%
50
45
40
35
30
25
20
15
10
5
0
Alnus undiff.
Aster type
Betula
Campanula
Caryophyllaceae
Corylus
Cyperaceae
Epilobium type
Ericales
Lactuceae
Lemna
Lycapodium undiff
Monolete undiff.
myriophyllum
Myrica
Picea
Pinus
Poaceae
Polygonum aviculare
Ranunculus Und
Rumex
Salix
Selaginella selaginoides
Trilete undiff.
Trifolia (e.f.)
unknown/unidentifiable
Urtica (e.f.)

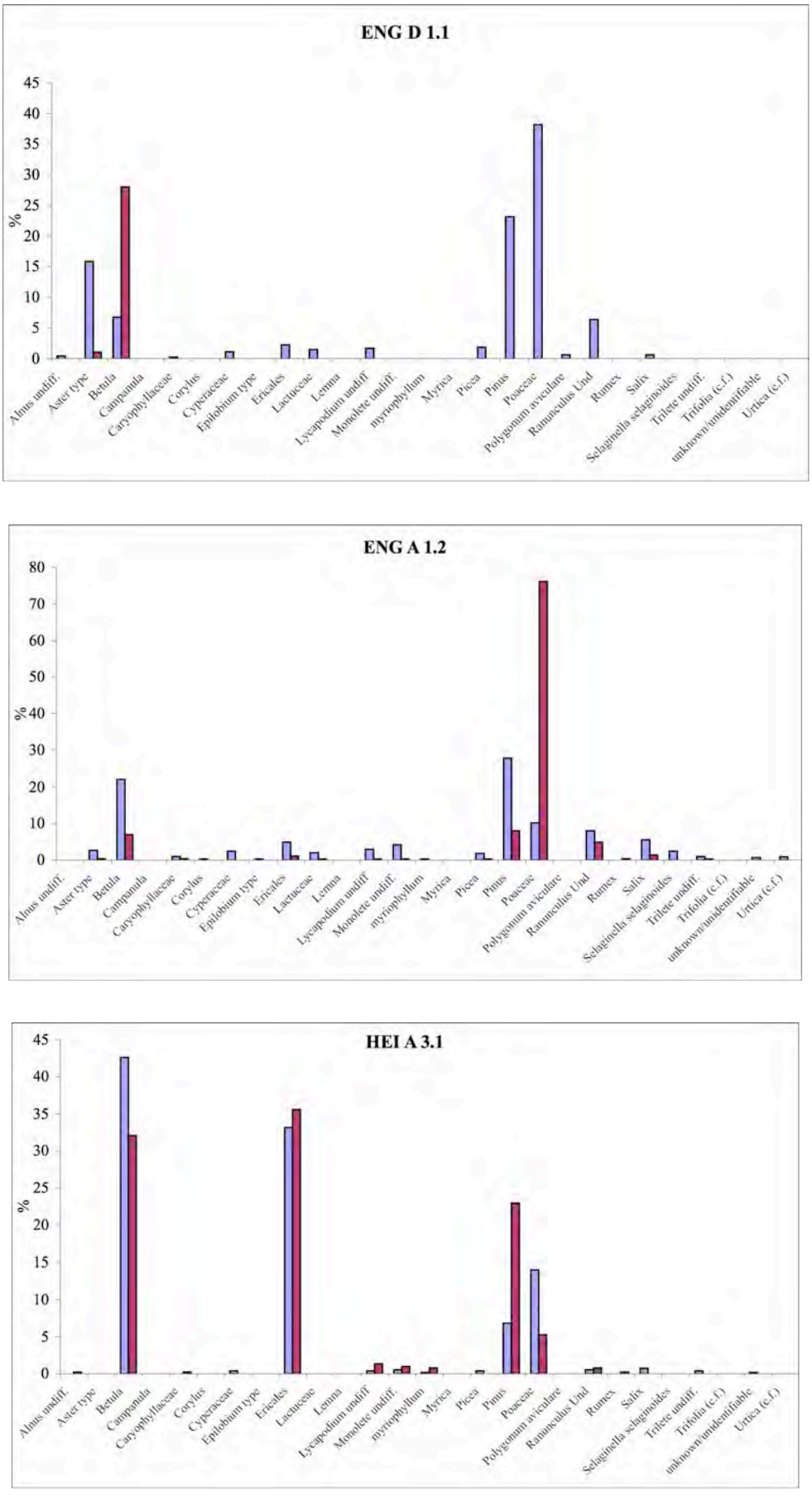


Figure 4.4 Percentage soil pollen (blue) and moss pollster pollen (red) bar charts.

Taphonomic processes appear not to have impoverished the pollen composition in soil samples relative to the moss. In fact the soil samples tend to be slightly more diverse than the moss samples (Table 4.1) which could be accounted for by the longer timescale which is likely represented.

HEI samples reflect predominantly locally derived heathland taxa with assemblages dominated by Heathers and Birch and some grass pollen. ENG samples are also predominantly locally derived with assemblages heavily dominated by grass pollen.

Interestingly pine occurs as a very similar proportion (~18%) of the assemblage in both HEI and ENG samples implying a relatively constant influx rate into both ecosystems and offering the possibility that pine might be usable as an indicator of the relative influx rate of other pollen types.

4.5.4 Comparison of Plant Floral Survey, Plant DNA and Plant Pollen data

As can be seen from the tables above, the three types of data give results with different biases and with different levels of resolution. Past plant assemblages are traditionally inferred from identification of fossil pollen or macrofossils, but taxonomic resolution is limited, particularly in the Arctic (Colinvaux 1964). Nevertheless the overlap between the plant data results and the DNA data is considerable; of the 72 plants botanically identified to species level 17 species had also been identified to species level by cpDNA and at the genus level, the figures are 48 and 23 respectively.

Typically the plant data and DNA data are able to identify plants to the species level, whereas the pollen data tends only to be able to resolve flora to lower taxonomic levels. Thus in order to evaluated the results from all three sources of information the data have been taken at the family level, and only the results from the same quadrants have been compared and plotted as area-proportional Venn diagrams (Figure 4.5).

DNA sequences were assigned to species level by conducting a local BLAST using the *trn*L database of 856 species of arctic plants (289 genera in 84 families) (Taberlet *et al. in prep.*).

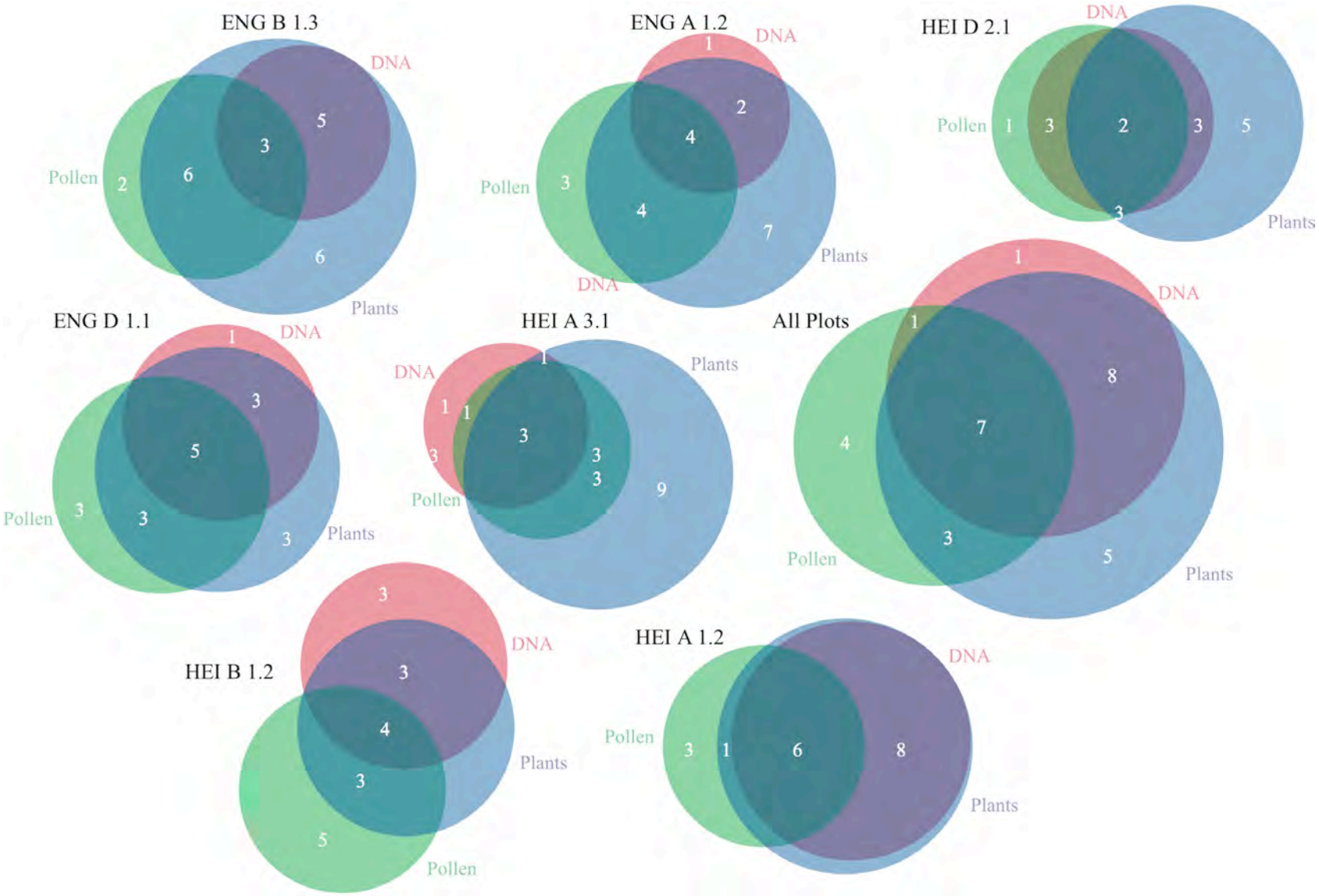


Figure 4.5 Area Proportional Venn diagrams of family counts from pollen, DNA, and plant data.

The resolution of the pollen data is generally less than that for the plant survey and DNA results, but when the three data types are compared at the family level, the degree of overlap is again considerable. This high level of coincidence is even more remarkable considering the plant data presented in the Venn diagrams were gathered in 50x50 m plots whereas the pollen and DNA data are derived from less than 10 g of sediment, and the DNA sequences derived from 1 µl of extract! This overlap would be even more pronounced if the Pinaceae family is excluded which would be valid as pine pollen is not of local origin, and no members of the Pinaceae family were documented in the plant data. The occurrence of *Pinus* sequences can be attributed to chloroplast DNA being recovered directly from pine pollen in the sediments. It has been previously reported that *Pinus sylvestris* pollen isolated from Holocene lake sediments yielded cpDNA when single grains were added to PCR mixes (Bennett and Parducci 2006). Although the sedimentary DNA extraction technique would have removed all pollen grains before the amplification step, it is probable that DNA from the pollen would have been extracted during this process.

The number of plant species identified through cpDNA would be further increased if different primers were used which can discriminate species within certain genera, such as *Salix*, and also if the number of 454 reads per sample were to be increased. This would further increase the correlation between the DNA data and the plant data at the species level.

For the above reasons is not surprising that the DNA species biodiversity should represent a sub-sample of plant diversity, but it is very interesting that species no longer present in the immediate environment have been uncovered. This could result

from the species being missed or misidentified during the floral survey, from sequencing errors giving rise to the misinterpretation of new species, or reflect species turnover in the environment with the sedimentary DNA reflecting an earlier period.

4.5.5 Ancient Sediments

The different plant haplotypes recovered from the permafrost samples are listed in Appendices 4 and 5, along with mammal species recovered from the same samples. These samples cover the period including the Last Glacial Maximum which spans the interval from 25 000 to 15 000 BP (Ray and Adams 2001) but peaked about 21 000 years ago (Otto-Bliesner et al. 2006). The plant species recovered are typical of a treeless vegetation dominated by arcto-boreal deciduous or evergreen low and high shrubs, sometimes with tussock-forming graminoids and true mosses, bog mosses, and lichens (Bigelow et al. 2003). The species recovered were assigned to bioclimatic zones using the PanArctic Flora Checklist (Elven 2007) and temperature zone interpreted using The Circumpolar Arctic Vegetation Map (Walker et al. 2005). See Table 4.3 for a description of the bioclimatic zones. A summary of findings is given below, with a more detailed breakdown in Appendix 4.

Table 4.3 showing Bioclimatic Zone for Circumpolar Arctic Plants and Temperature Conditions

Bioclimatic Zone	**Description of Zone**
A	Polar desert, mean July temperature 1 to 3 ºC
B	Northern Arctic Tundra, mean July temperature 4 to 5 ºC
C	Middle Arctic Tundra, mean July temperature 6 to 7 ºC
D	Southern Arctic Tundra, mean July temperature 8 to 9 ºC
E	Shrub Tundra, mean July temperature 10 to12 ºC
N	Boreal and/or Boreal-Alpine

N.B. "N" is not a well defined bioclimatic zone. It simply refers to boreal and alpine areas neighbouring the arctic.

EH336: 34085±243 calBP

Most of the taxa from this sample are found in bioclimatic zone E (shrub tundra) and/or are rare in zone D (*e.g. Avenella flexuosa, Trientalis europaea, Trollius europaeus, Anthoxanthum nipponicum, Anemone patens=Pulsatilla patens*). The lack of tree species (except Salicaceae) indicates open vegetation.

EH333: 32617±311 calBP

The bioclimatic zone suggested by plant species from this sample is E or N based on the occurrence of *Stellaria borealis* and *Picea*. However, as only five taxa were informative, the resolution is not high. The conditions were probably not as of bioclimatic zone A or B because none of the taxa occur in zone A and only one rarely occurs in zone B.

EH287: 30275±305 calBP

Most species found are frequent in bioclimatic zone E, although temperatures could be warmer than typical for this zone as many species are also are frequent in zone N (boreal or boreal-alpine). However, the occurrence of *Lagotis glauca* and *Ranunculus arcticus*, which only are scattered in zone N, point towards zone E. The coldest possible conditions are zone D as some species are frequent in this zone but the occurrence of *Pulsatilla patens* and *Crepis chrysanta*, which are rare in zone D, indicates that the conditions were somewhat warmer. Conditions found in climatic zone C is unlikely due to the occurrence of *Angelica sylvestris/Cnidium cnidiifolium, Pulsatilla patens, Stellaria borealis, Crepis chrysantha, Trisetum sibiricum* and *Betula.*

EH316: 27553±433 calBP

The most likely conditions are as in bioclimatic zone E due to the lack of tree species (except Salicaceae) and owing to the occurrence of some species which are only scattered or are rare in zone N (*Armeria scabra* and *Lagotis glauca*). The coldest possible conditions are as in zone D but this is not likely as so many species are rare in this zone (*Thymus oxyodontus, Avenella flexuosa, Trientalis europaea, Anthoxanthum nipponicum, Crepis chrysantha, Trollius europaeus, Sanguisorba officinalis, Anemone patens=Pulsatilla patens,* and *Trisetum sibiricum*).

EH308: 24816±221 calBP

The species recovered are suggestive of bioclimatic zone N. *Alnus incana* and *A. viridis* indicate pioneer boreal forest.

EH288: 19026±224 calBP

The occurrence of *Larix* indicates boreal forest (bioclimatic zone N). *Larix* is characteristic of a forest-dominated environment, containing boreal broad-leaved or needle-leaved deciduous trees, sometimes with significant component of arcto-boreal evergreen needle-leaved trees. The most commonly recovered sequence was of *Trollius europaeus* which only occurs from one other sample, and there in low abundance. Although an arctic-alpine plant (Pompanon et al. 2006), this species is characteristic of moist meadows, damp places in grasslands or woods (Czerepanov 1995) its USDA Plant Hardiness rating of 5-9 equates to an average annual minimum temperature of –28.9 °C to –1.1 °C (Cathey 1990). All the other species, such

Gentianopsis are also boreal and low arctic species strongly suggesting that boreal conditions are prevailing.

The bioclimatic zone designations given to the plant communities recovered from the permafrost sediment samples corresponds to the insect data (Figure 4.6) (particularly the relative abundance of arctic insects), and also to the marine isotope curve (Figure 4.7). The exception is EH308, which is designated as bioclimatic zone N, yet is dated to 24816±221 calBP, close to the peak of the Last Glacial Maximum (Otto-Bliesner *et al.* 2006). The presence of zone N species may therefore indicate a localised or brief warming event, and this interpretation is again supported by the absence of arctic insects at this time.

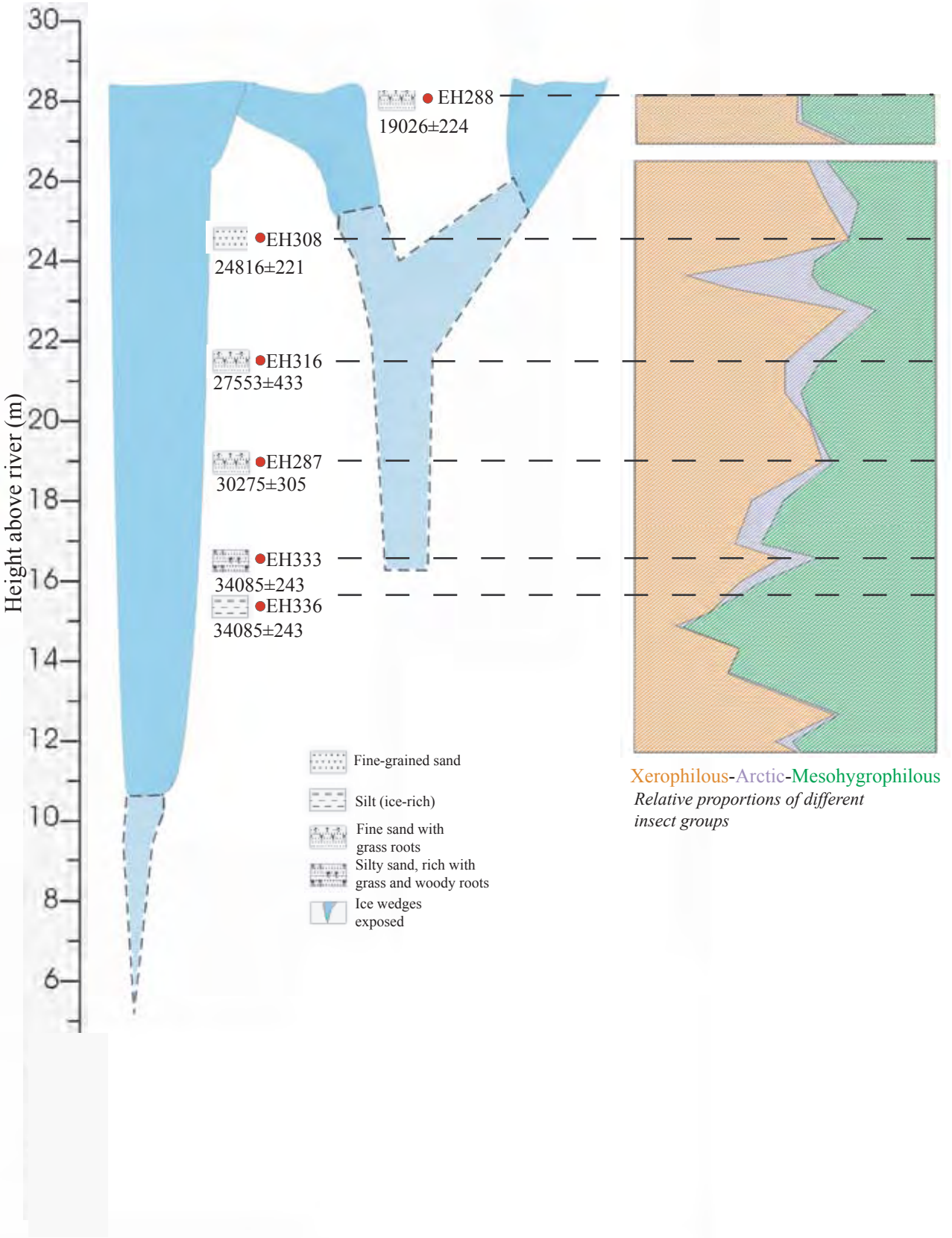


Figure 4.6 Vertical section of Ice Bluff Cliff showing sediment sampling sites and dates (calBP), compared to relative proportions of xerophilous, arctic and mesohygrophilous insect groups (Kuzmina unpub. data).

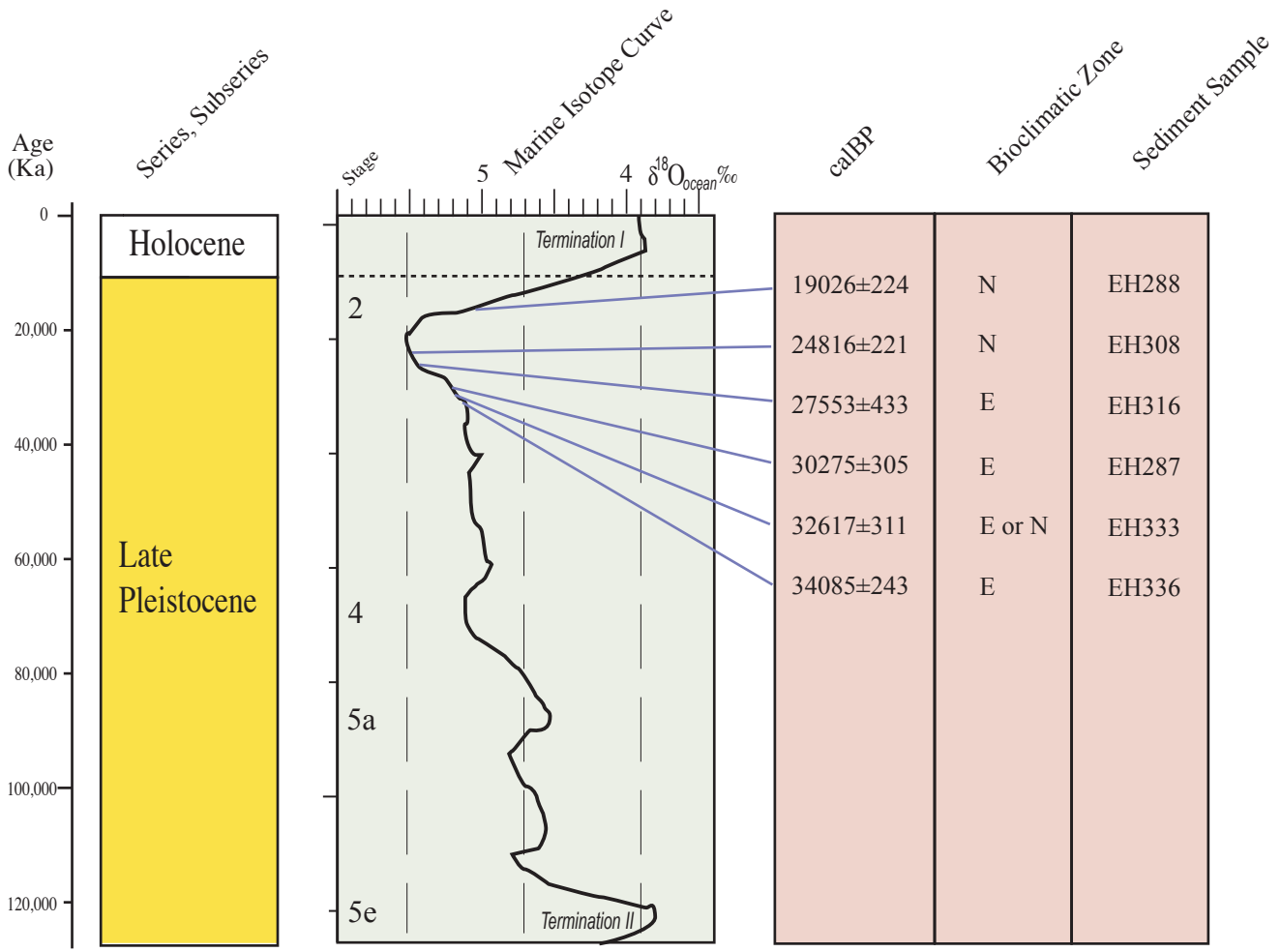


Figure 4.7 Showing marine isotope curve against bioclimatic zone. Modified from Gibbard, P.L. and colleagues (2005).

4.6 Conclusions

Although it is theoretically possible from pollen morphology to identify all the families present at Varanger, some families (*e.g.* Ericales) are difficult to distinguish using fossil pollen because of poor preservation or indistinct morphological characters. Furthermore the resolution of pollen at lower taxonomic level varies considerably among families. It has been estimated that the maximum resolution based on well-preserved fossil pollen resulted in identification of 46.5% of the genera and only 10.2% of the species (Sønstebø et al. in prep.). In the major arctic families of Cyperaceae, Poaceae, Salicaceae, Brassicaceae and Caryophyllaceae, limited resolution can be achieved from fossil pollen, and species-level resolution is often difficult in the Saxifragaceae.

Additionally, the low pollen productivity of insect-pollinated plants tends to result in poor representation of many species in the fossil record. Prolific wind-pollinated species (particularly *Betula* and *Alnus*) are over-represented and further skew pollen assemblage composition.

However, the DNA method, whilst allowing higher taxonomic identifications also has limitations. For example, some important genera (*e.g. Artemisia* and *Populus*) are identifiable based on pollen, but not by *trn*L*g/h* region of the cpDNA. This can be overcome by targeting a different area of the chloroplast genome, species in these genera would become distinguishable.

Having established the relationship between pollen data, plant data, and DNA data, this DNA-biodiversity assessment technique can now be securely applied to investigations into past species' distributions and previous species assemblages. It will help to refine already established scenarios of species turnover such as the invasion of trees during the Holocene (van der Knaap et al. 2005), and allow testing of key assumptions of models, such as niche stability and the persistence of species-environment relationships and species assemblages over time and space.

Chapter 5

Megafaunal response to Catastrophic Change

No moa, no moa
In old Ao–tea–roa.
Can't get 'em.
They've et 'em;
They've gone and there ain't no moa!

Traditional

5.1 Abstract

In recent years several studies have reported the successful extraction of ancient DNA (aDNA) from both frozen and non-frozen sediments (even in the absence of macrofossils) in order to obtain genetic "profiles" from past environments. One of the hazards associated with this approach, particularly in non-frozen environments, is the potential for vertical migration of ancient DNA across strata. To assess the extent of this problem, we extracted aDNA from sediments up to 3300 years old at two cave sites in the North Island of New Zealand. These sites are ideal for this purpose, as the presence or absence of DNA from non-indigenous fauna (such as sheep) in sediments deposited prior to European settlement can serve as an indicator of DNA movement. Additionally, these strata are well defined and dated. DNA from sheep was found in strata that also contained moa DNA, indicating that genetic material had migrated downwards. Quantitative PCR analyses demonstrated that the amount of sheep DNA decreased as the age of sediments increased. Our results suggest that sedimentary aDNA is unlikely to be deposited from wind-borne DNA, and that physical remains of organisms, or their ejecta, need to have been incorporated in the sediments for their DNA to be detected. Our study indicates that DNA from sediments can still offer a rich source of information on past environments, provided that the risk from vertical migration can be taken into account.

5.2 Introduction

Ancient DNA (aDNA) from diverse mammals and plants has been obtained directly from minor amounts of permafrost (permanently frozen) sediments many thousands of years old (Willerslev *et al.* 2003; Lydolph *et al.* 2005). Likewise, under *non*-frozen conditions, trace amounts of sediment have yielded aDNA sequences of diverse vertebrate and plant species, even in the absence of macrofossils (Hofreiter *et al.* 2003; Willerslev *et al.* 2003). The immediate origin(s) of this DNA, is unclear. A possible source of plant DNA in sedimentary deposits is fine rootlets (Willerslev *et al.* 2003). A variety of sources have been suggested for animal DNA, including dung, urine, skin, hair, and keratin (Lydolph *et al.* 2005). Also uncertain is whether the DNA is extracellular and bound to clay minerals, or if cellular DNA is released during the extraction procedure (Ogram *et al.* 1988).

To date, most aDNA sedimentary analyses have examined soil profiles from permafrost regions (Hansen *et al.* 2001; Willerslev *et al.* 2003; Willerslev *et al.* 2004; Willerslev *et al.* 2004; Lydolph *et al.* 2005; Mitchell *et al.* 2005). In this environment, two sources of evidence suggest that DNA leaching and re-deposition are not significant problems: firstly, changes observed in floral and faunal communities through time agree broadly with those predicted by macrofossil records (Willerslev *et al.* 2003; Lydolph *et al.* 2005); secondly, the recovered DNA fragments were damaged in clear age-dependent patterns, despite discontinuous sediment chronology and the presence of free water (Hansen *et al.* 2001; Willerslev *et al.* 2004). These results are encouraging, but the potential for DNA being leached in non-frozen conditions remains to be examined (Pääbo *et al.* 2004). This is especially important because temperate and desert cave sites are major sources of aDNA used in

reconstructing past environments (Poinar *et al.* 1996; Poinar *et al.* 1998; Hofreiter *et al.* 2000; Hofreiter *et al.* 2003; Willerslev *et al.* 2003). DNA leaching would significantly complicate, or even invalidate, the interpretation of results in some contexts (Poinar *et al.* 1996); (Poinar *et al.* 1998; Hofreiter *et al.* 2000). If the fidelity of DNA sequences can be established from strata (of defined age), however, then this technology will enable paleofaunal reconstructions spanning thousands of years (Willerslev and Cooper 2005). Apart from leaching, other potential sources of bias in taxon representation include taxon-dependent factors such as body size, and the likely presence of remains such as hair, feathers, eggs, and dung.

In this study, we used sediments from two cave sites in New Zealand to gain insights to the origin of sedimentary aDNA and to test for DNA leaching in non-frozen sediments. New Zealand has an environment ideal for investigating possible DNA leaching because the preservation of bones and other materials is excellent. Most importantly, New Zealand had a limited range of large vertebrates, mostly birds, prior to the arrival of humans; terrestrial mammals were entirely absent, with the exception of three bat species. This distinctive faunal composition allows leaching to be readily identified, because the presence of non-indigenous mammal species in pre-settlement strata is necessarily the result of downward movement of DNA in the sediments. For example, sheep (*Ovis aries*), whose numbers currently exceed 40 million, were introduced to New Zealand by European settlers only from the 1830s; they did not reach the study area until c. 1870. It is also well established that the large, indigenous, herbivorous ratite birds known as moa (Aves; Dinornithiformes) became extinct by c. 550 years ago (Holdaway and Jacomb 2000), and so the presence of moa DNA in European layers would imply upward movement of sedimentary DNA or of the

sediment constituents that retained the DNA. These and other distinct changes in the biota provide an opportunity to assess the extent of DNA leaching within non-permafrost sediments.

5.3 Materials and Methods

Sediment samples were taken from freshly excavated sections in two dry caves located on the southern face of the Hukanui range, North Island, New Zealand (Figure 5.1): Hukanui Pool and Hukanui #7a (Figure 5.2), are c. 300m apart, and at 860 and 800m asl, respectively (Holdaway and Beavan-Athfield 1999). The sites are beneath large erratic limestone blocks and both contain sediment layers ranging in age from >3000 years old to the present. Importantly, the Hukanui locality itself is well-suited for investigating paleogenetic reconstructions because it contains two layers of volcanic ejecta from massive eruptions from the Lake Taupo volcano (Figure 5.1); the Taupo ignimbrite deposited at AD 232 ±15 (Sparks *et al.* 1995) and the Waimihia tephra, deposited at 3280 ± 20 yr ^{14}C years before present; (Froggatt and Lowe 1990).

The Taupo eruption emplaced the Taupo ignimbrite from a pyroclastic flow, which travelled the 80 km to the sites at high speed and at a relatively high temperature, destroying all vegetation and fauna in its path. It disturbed the sediment surfaces and entrained bones and sediment particles in the sites, and now forms a tephra layer up to 600mm thick, with charcoal, sediment, and (rare) bone inclusions. The Taupo ignimbrite has the consistency of bulk cement and can also support vertical sections; all disturbances in the layers are easily visible.

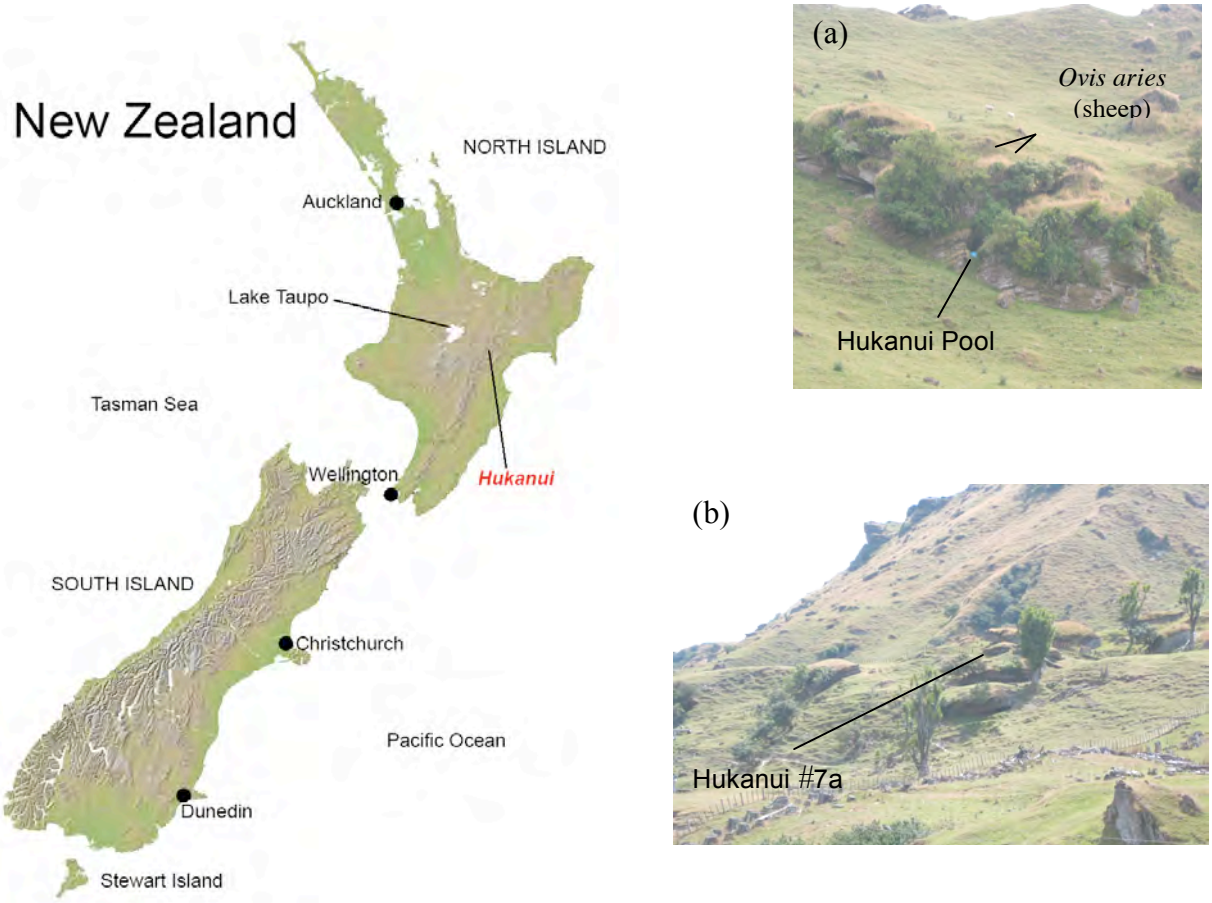


Figure 5.1 Map of New Zealand, showing location of Hukanui sites: (a), General view of rock pile containing Hukanui Pool site; (b), General of view of rock pile containing Hukanui #7a site.

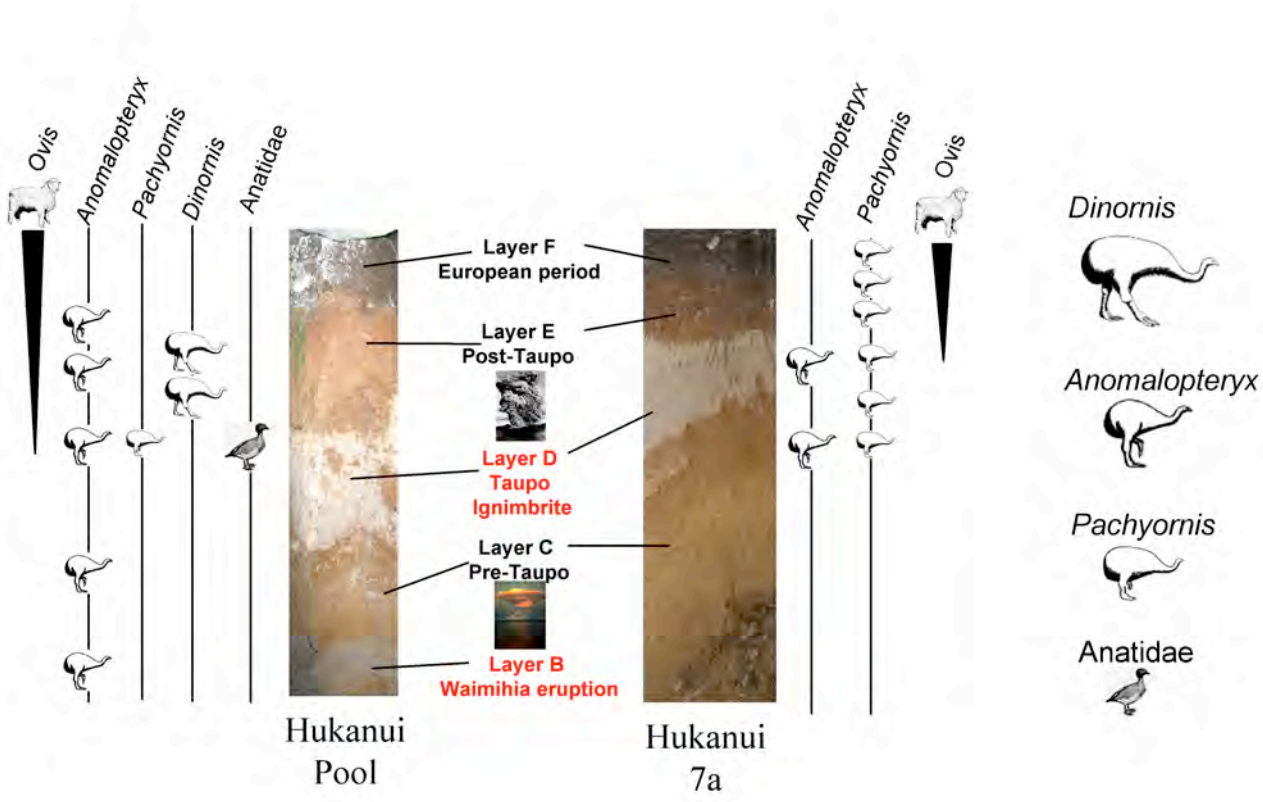


Figure 5.2 Stratigraphic sections in Hukanui Pool and Hukanui #7a sites showing sampling sites, and strata. The Hukanui #7a section is viewed into a corner of an excavation square, and the slope of the interfaces is exaggerated. Note the irregular interface at the base of the Taupo Ignimbrite, reflecting the disturbance of the sediment surface by the incoming tephra.

By contrast, the Waimihia tephra (Froggatt and Lowe 1990) was deposited as a "rain" of pumice lapilli c. 1-2 mm in diameter, which stripped leaves from trees but did not necessarily kill them. The tephra formed a layer c. 100 mm thick (Froggatt and Lowe 1990) in the Hukanui Pool site, which was more open, but did not enter Hukanui #7a in any significant amounts.

The sediments in Hukanui Pool are more or less horizontal over most of the floor, but the layers slope downwards towards the western wall. In Hukanui #7a, sediments were incorporated both as horizontal fill amongst and above boulder debris and as a small fan in the north-western corner of the excavation, where the aDNA samples were collected. The fan had been formed from sediments entering through a small opening from an upper chamber (Holdaway *et al.* 2002).

Additionally, the sharp, sand-sized particles of fossil barnacle shells and other marine organisms (derived from the enclosing limestone rocks) that compose much of the cave sediments contain variable amounts of clay. The clay binds the sediments, making them firm and cohesive and lending stability to vertical sections, so that macroscopic particles cannot migrate through or between the layers. Any DNA found out of stratigraphic context must therefore have migrated as free molecules, or bound to microscopic particles. Significantly, there are no rooted plants nor soil horizon development in the sediments of either site, so they are not soils.

The sediments were moist as a result of condensation in a humid environment, but there is no fall of water onto the surface, nor flow across it. These are rock shelters, not true water-carved caves and there is no free water flowing in the sites (Holdaway

and Beavan-Athfield 1999). Both sites are just a few meters from the open air and the sediments grade continuously into the outside soil.

5.4 Sampling, DNA extraction, and amplification.

Contamination with extraneous DNA is an ever-present concern in any aDNA study and it is the responsibility of the researcher to demonstrate that adequate experimental and authentication procedures are carried out (Cooper and Poinar 2000). Samples were taken from freshly excavated sections in the two shelters, beginning at the bottom of each section and proceeding to more recent levels (Figure 5.2). Disposable tools were used and changed between samples to avoid cross-contamination. All manipulation of ancient samples before polymerase chain reaction (PCR) amplification were performed in dedicated aDNA laboratories at the Henry Wellcome Ancient Biomolecules Centre, University of Oxford, and the Centre for Ancient Genetics at the University of Copenhagen, in areas free from other molecular research. One negative extraction and one amplification control was used for every eight samples extracted and each positive result cloned a minimum of eight times in agreement with suggested aDNA criteria (Handt *et al.* 1994; Willerslev *et al.* 1999; Hansen *et al.* 2001; Gilbert *et al.* 2003; Willerslev and Cooper 2005).

DNA from a total of 1g of wet weight sediment per sample was extracted in two sub-samples of 0.5g sediment, dissolved in 600µl lysis buffer (Bulat *et al.* 2000) 400 ug/ml proteinase K (Roche) disrupted with four runs of a FASTprep 120 (BIO 101) at speed 6.5 for 45 sec, with 2 min on ice between runs and incubated at 65 °C for at least 4 h under agitation. The solution was adjusted with NaCl to 1.15M, treated with 1/2 vol. of chloroform/octanol (24:1), and agitated slowly overnight at room

temperature, and the water phase isolated with centrifugation at 12 000 g for 2 min and transferred to a separate microtube for incubation at 2-3 °C for at least 1 h. The precipitate was centrifuged at 12 000 g for 2 min and the supernatant purified using silica spin columns and PB-buffer (Qiagen DNA purification kit II), followed by washes in 0.5 ml Salton wash 1 and 2 (BIO 101), and 0.5 ml AW1 and AW2 (Qiagen tissue kit). The DNA was eluted twice with 100 µl EB buffer (Qiagen purification kit II) and stored at –20 °C.

PCR was used to amplify an 88 bp (moa) and 60 bp (sheep) fragment of control region mitochondrial DNA (mtDNA), avian 153 bp fragment of 12S mtDNA, and plant *rbcL* and *trnL* chloroplast DNA using primers listed in Table 5.1, using 5 µl of DNA extractions, 35-55 cycles of PCR (1.5 min initial denaturation at 94 °C, 45 s at 94 °C, 45 s at 45-60 °C, 1.5 min at 68 °C and a final cycle of 10 min at 68 °C). PCR products were cleaned using a QIAquick PCR Purification Kit (QIAGEN). Amplification products from the two separate extracts of each sample were pooled, cloned, purified, and sequenced on both strands (Willerslev *et al.* 1999). Sequences were aligned using ClustalW in BioEdit (Hall 1999) and possible recombination among the clone sequences investigated (Willerslev *et al.* 1999).

Table 5.1 PCR Primer names and sequences are shown, along with references, annealing temperatures, and approximate product length.

Name	**Sequence**	**Reference**	**Annealing temp. (°C)**	**Product length (bp)**
Moa CR 262F	5'-GCGAAGACTGACTAGAAGC	(Cooper *et al.* 2001)	55	88
Moa CR 329R	5'-ATACGAGAAATAGGGATTAAGC	(Cooper *et al.* 2001)		
12SE	5'-CCCACCTAGAGGAGCCTGTTC	(Cooper *et al.* 1992)	52	153
12SH	5'-CCTTGACCTGTCTTGTTAGC	(Cooper *et al.* 1992)		
rbcL h1aF	5'-GGCAGCATTCCGAGTAACTCCTC	(Poinar *et al.* 1998)	45	130
rbcL h2aR	5'-CGTCCTTTGTAACGATCAAG	(Poinar *et al.* 1998)		
*Trn*L*g*	5'-GGGCAATCCTGAGCCAA	n/a	60	87
*Trn*L*gh*	5'-CCATTGAGTCTCTGCACCTATC	n/a		
Upper Carex 18-mer	5'-GCAATCCTGAGCCAAATC	(Taberlet *et al.* 2007)	55	88
Lower Carex 18-mer	5'-CCATCGAGTCTCTGCACC	(Taberlet *et al.* 2007)		
Ovis CR 1F	5'-CATAATGGTAAGCATGGGCAT	n/a	60	60
Ovis CR 1R	5'-CACGATACAGCAGATATGTC	n/a		
Sheep-87F	5'-CCTGTCCATTAGATCACGAGCTTG	n/a	60	71
Sheep-157R	5'-GAGGGATCCTTGCTAAGCGG	n/a		

5.5 Quantitative Real-Time PCR (qPCR)

A SYBRGreen based qPCR assay was used to determine the relative quantity of sheep DNA within the Hukanui Pool DNA extracts. Amplifications targeted a 71 bp fragment of the sheep mtDNA control region, using qPCR primers Sheep-87F and Sheep-157R (Table 5.1). Before the qPCR analysis, the primers were pre-screened on both sheep DNA-positive and -negative soil extracts and blanks, to ensure that they generated a single correct product, with no primer-dimer or non-specific products that might contribute to erroneous results. The qPCR assay itself was performed in 25 μl reaction volumes, using dilution series of the Hukanui Pool extracts (1 μl, 0.5 μl, 0.25 μl, 0.125 μl concentration) to ensure replicability of the results and to screen for any inhibitory effects that might complicate interpretation of the results. The reaction contained 300 μM (final concentration) of each primer, used SYBRGreen PCR Mastermix (Applied Biosystems), and was performed using an ABI PRISM 7000 instrument (ABI). A dissociation curve was generated at the end of the reaction, to ensure that measured fluorescence was attributable only to the correct amplification product (through assessment of fragment melting temperature). As a further control the amplified products were analysed using conventional agarose gel electrophoresis to ensure the correct size of the amplified product (Table 5.4).

5.6 Sequence Identification

Sequences were assigned to species using a Bayesian approach that allows calculation of probabilities of membership of specific taxonomic groups. This automated procedure combines database searches with alignment algorithms and Bayesian phylogenetic procedures, using information from NCBI's taxonomy browser (Benson *et al.* 2000; Wheeler *et al.* 2000). First, using database searches we identified sets of closely related sequences to each query sequence. The 50 best hits defined by E-values of a BLAST search were chosen, eliminating identical copies and requiring at least three families and two orders to be sampled for each query sequence. An alignment of the 50 non-redundant sequences was then produced using ClustalW (Thompson et al., 1994). Flanking regions up to 200 bp were included, if available, from the database sequences.

The sequences were then analysed using MrBayes (Hulsenbeck and Ronquist 2001), with the GTR model of nucleotide substitution. In the Markov chain Monte Carlo (MCMC) analysis, two simultaneous chains were used with 1 000 000 steps and a burn-in of 100 000 steps. k = 10 000 trees were sampled from the results and, based on the taxonomic assignments of sequences in Genbank, the probability of each query sequence forming a monophyletic group with sequences from a particular phylogenetic group was estimated from the posterior sample of trees. For example, the probability that a query sequence (Q) belongs to family F would be assessed as

$$\Pr(Q \in F \mid X) = \int_{G \in \Omega} I(Q, F \text{ monophyletic in } G)\, p(G \mid X) dG \approx \sum_{i=1}^{k} I(Q, F \text{ monophyletic in } G_i)$$

where □ is the set of all possible trees (G) with branch lengths, G_i is the ith tree sampled from the MCMC in MrBayes, X is all of the sequence data, and is an indicator returning 1 if Q and F are monophyletic in tree G and 0 otherwise. In general, we did not require all sequence belonging to F to form a monophyletic group as long as some sequence in F and Q form a monophyletic group. This inference procedure was performed for each sequence independently for all possible taxonomic assignments from the level of class to the level of species. There are several limitations of this method, the most important being that assignment can only be done to taxonomic groups represented in the database. If only one relevant species is represented in the database, the query sequence will be assigned to this species with a probability of one.

In the vast majority of cases, the query sequences could be assigned to particular species with very high posterior probability ($PP > 0.9$).

5.7 Results and Discussion

5.7.1 Avian DNA

Avian DNA was successfully amplified from the layers shown in Figure 5.2 and listed in Table 5.2. The absence of moa DNA from Layer F in Hukanui Pool, the youngest layer in that site, is expected because moa have been extinct in New Zealand for at least 500 years and this result confirms that upward movement of the DNA in this site is not a detectable problem. The presence of amplifiable moa DNA in Layer F in Hukanui 7a is not surprising, however, because the layer had been affected by an earlier (1959) excavation that reached the top of the Taupo ignimbrite (Layer D) over part of the site, and sediment from pre-European levels elsewhere in the shelter had been displaced and mixed with the European layer materials at the sampling point. In addition, the sample section was beneath the opening to an upper chamber in the site complex (Holdaway et al. 2002), through which sediment and other materials have entered. At the lowest, oldest stratigraphic levels, amplifiable moa DNA was absent from Layer A in Hukanui Pool, and Layers A and B in Hukanui 7a; this could be the result of either degradation or limited levels of occupation at that time in each site, or both.

The absence of moa DNA from the youngest and oldest intact layers, and the restriction of certain moa genera to specific strata (Table 5.2), suggest that vertical movement of avian DNA is not a significant factor in these sites, thereby permitting an environmental interpretation of the results.

The pattern of occurrence of moa taxa in Layers B-E of the two sites (Table 5.2) appears to be related to the physical constraints of both study sites (Figure 5.3), the

vegetation surrounding the sites, and possibly to interactions between the moa taxa that used the sites for nesting. The lower chamber of Hukanui #7a has always had a very low ceiling (Holdaway *et al.* 2002), and bones (Worthy and Holdaway 2000) and DNA (Table 5.2) of only the two smaller taxa were found in the sediments. However, although their macroscopic remains and DNA have not been identified yet in the deposit, very small juvenile *Dinornis* could have entered both the upper and lower chambers. Although Hukanui Pool was always large enough for all three taxa to enter (Figure 5.3), only the two larger taxa are represented there by macrofossils (Holdaway unpublished data). DNA of all three taxa was identified from Hukanui Pool (Table 5.2), but DNA of the smallest moa (*Pachyornis*) was recorded only from near the surface of Layer D (Taupo ignimbrite). The apparent absence, except for one brief period, of *Pachyornis* from Hukanui Pool, which it physically could have entered (Figure 5.3), may be explained by factors such as habitat preference, environmental disturbance resulting from the most recent Taupo eruption, and interaction between the species.

Table 5.2 Stratigraphic distribution of moa and other bird taxa in Hukanui Pool and Hukanui #7a.

*Numbers in brackets represent the highest posterior probabilities over obtained for that genus (or for Anatidae, for that family). Layer numbers represent distance in mm below strata interface (*e.g.*, Layer B-10 refers to sample in Layer B 10mm below B/C boundary). *Independently produced results.*

		Hukanui Pool		
		Primers		**Primers**
Strata	**Sample No.**	**Moa CR 262F, 329R**	**Sample No.**	**12SE, 12SH**
F European	Layer F	No products		No products
E Post-Taupo	Layer E-25	*Anomalopteryx* (93%)		*Anomalopteryx* (100%)
E Post-Taupo	Layer E-150	*Anomalopteryx* (97%) *Dinornis* (79%)	Layer E-150	*Anomalopteryx* (97%)
E Post-Taupo	Layer E-290	*Dinornis* (81%)*		
D Taupo	Layer D-20	*Anomalopteryx* (98%) *Pachyornis* (99%) *Pachyornis* (100%)*	Layer D-20	*Anomalopteryx* (98%) Anatidae (100%)
C Pre-Taupo	Layer C-10	No products	Layer C-10	*Anomalopteryx* (97%)
B Waimihia	Layer B-10	*Anomalopteryx* (92%)	Layer B-10	*Dinornis* (52%)
A Pre-Waimihia		No products		No products

	Hukanui #7a			
		Primers		**Primers**
Strata	**Sample No.**	**Moa CR 262F, 329R**	**Sample No.**	**12SE, 12SH**
F European	**Layer F**	*Pachyornis* (100%)*		**No products**
E Post-Taupo	**Layer E-10**	*Pachyornis* (100%)*	**Layer E-10**	No products
E Post-Taupo	**Layer E-80**	*Pachyornis* (91%)	**Layer E-80**	No products
D Taupo	Layer D-100	*Pachyornis* (100%)	Layer D-80	*Anomalopteryx* (96%)
C Pre-Taupo	Layer C-10 Layer C-80	*Pachyornis* (95%) *Pachyornis* (100%)	Layer C-10 Layer C-80	*Anomalopteryx* (98%)
B Waimihia	**Not present**		**Not present**	
A Pre-Waimihia	Not present		Not present	

The Taupo ignimbrite completely devegetated the slopes around the sites during its emplacement (Worthy & Holdaway 2000). The DNA evidence suggests that *Pachyornis* used the Hukanui Pool site only for a brief period following this eruption, when the surrounding vegetation would have been recovering via stages including seral shrubland. If shrubland did not provide suitable habitat for the two larger moa taxa, then they would not have been present to exclude the small *Pachyornis*. Consequently, *Pachyornis* could have occupied the site (as it continued to do downslope in Hukanui #7a) until the forest developed sufficiently for the larger taxa to recolonize the area. If the larger moa had indeed excluded *Pachyornis* from Hukanui Pool, it would imply competition for secure nest sites based on body size. The large (c. 1.5-2 kg) extinct harrier (*Circus* sp.), whose own bones and prey remains were present in both sites (Worthy and Holdaway 2000), could well have preyed on the smaller moa and led to the latter's preference for enclosed nest sites. There is some evidence for harrier predation on small moa in the form of characteristic damage to at least one pelvis from Hukanui #7a (Holdaway unpublished data).

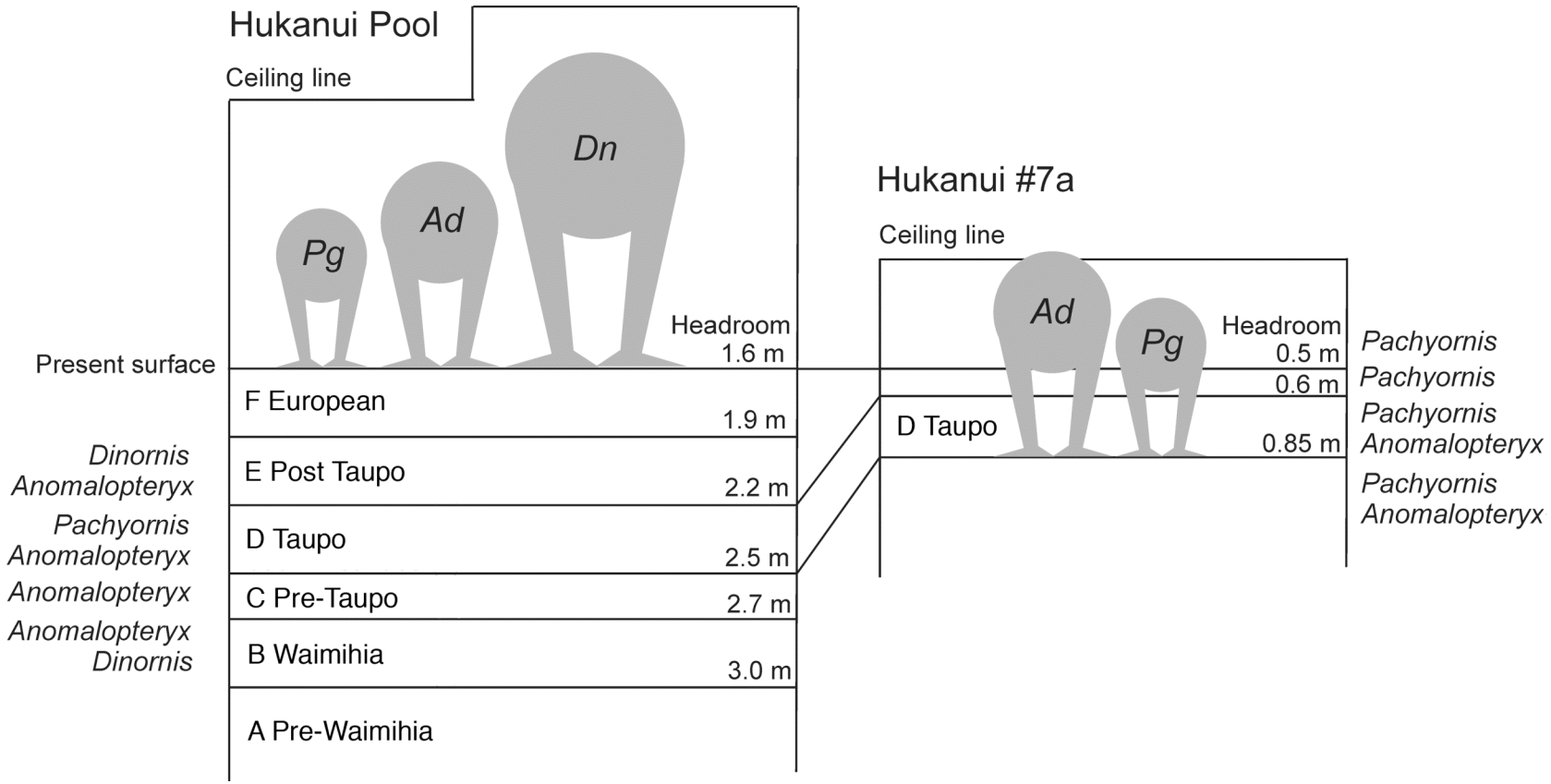


Figure 5.3 Relative size of moa species and the different sediment levels at both sites. *Pg* Pachyornis, *Ad* Anomalopteryx, *Dn* Dinornis.

The absence of Dinornis DNA from all sediments in Hukanui #7a, a site from which its larger juveniles and adults would have been physically excluded (Figure 5.3), and of *Pachyornis* from all but one level in Hukanui Pool, even though both species were present in the area, suggests that the sites record for the most part only those taxa that either physically entered the sites, alive or dead. Occasional deaths of very young *Dinornis* in Hukanui #7a could complicate the record, but do not appear to have done so. Hence, interpretation of presence-absence data based on DNA preserved in sediments has to take into account the physical and biological contexts.

Four species of moa have been identified from skeletal remains in the area of the sites: *Dinornis* (80-120 kg live mass), *Euryapteryx gravis* (40-70 kg), *Anomalopteryx didiformis* (30-40 kg), and *Pachyornis geranoides* (10-15 kg). Two of these genera (*Dinornis* and *Anomalopteryx*) are believed to have been associated with tall forest, although *Dinornis* was found also in areas with more open vegetation (Worthy and Holdaway 1993; Worthy and Swabey 2002). *P. geranoides* probably preferred wetland vegetation (Worthy & Holdaway 2002: 196), the shrubland ecotones provided by wetlands or the productive forest edge (Worthy 1990). There were no wetlands near the two sites, so the species seems also to have been able to occupy forest with an understorey of smaller angiosperms, which was the normal vegetation around the sites during the Holocene (Worthy and Holdaway 2000).

The mtDNA of *Pachyornis* identified in a sample (E-10) taken 10 mm below the "pre-European" surface in Hukanui #7a, and of *Anomalopteryx* in a sample taken 25 mm below the "pre-European surface" in Hukanui Pool (sample E-25), are the stratigraphically highest occurrences of DNA of any extinct taxa in intact sediments in

the two sites. Allowing for some disturbance of the floor of the shelter by the birds themselves, these occurrences represent the last stages of the presence of moa before their extinction in the area. Some deflation or erosion of the surface in each site may have occurred after deposition of the E-10 and E-25 material, but it is also possible that the sedimentation regime altered with the deforestation that accompanied Polynesian settlement of the area, at about the time of moa extinction. As noted above, the presence of *Pachyornis* DNA in Layer F of Hukanui #7a can be attributed to the mixing of debris from older layers with post-European debris during earlier excavations and there is direct evidence that moa material, and sediment originally emplaced in lower strata were mechanically transported by burrowing.

The lack of amplifiable avian DNA from the deep European layer in Hukanui Pool, where no previous excavations had disturbed the main deposit, indicates that birds have not inhabited the rock shelters in modern times, perhaps with the exception of occasional visits by foraging songbirds. It is perhaps more surprising that only two non-moa avian DNA sequences have amplified from New Zealand sediments, that of ducks (Anatidae, Hukanui Pool) and a parakeet at another site (Willerslev *et al.* 2003). However, the recovery of significant amounts of moa eggshells and bones from the Hukanui sites indicate that these were used by moa for nesting, so it is possible that the sheer volume of moa DNA has swamped that of any other species. Bones of non-moa taxa were rare and thinly distributed in Hukanui Pool (Holdaway unpublished data). Hukanui #7a, however, contained a rich deposit of bones (Worthy and Holdaway 2000; Holdaway *et al.* 2002) but these were concentrated 1-2 m from where the DNA sediment samples were taken later. The data indicate that relative biomass may need to be taken into account when using non-specific primers to

retrieve different elements of past ecosystems. Alternatively, it may be necessary to employ primers that preferentially amplify less common DNA. The Hukanui #7a results suggest also that more than one chrono-series of samples may be required to explore fully the diversity of DNA preserved in even a small site.

The Waimihia tephra is a volcanic airfall deposit, so it is not surprising that no DNA was retrieved from it. The presence of moa DNA in the Taupo ignimbrite is surprising, but could result from moa trampling faeces or feathers into the ash. However, given the particularly violent way the ignimbrite was deposited, it is likely that the DNA was attached to loose floor sediment that was torn up and incorporated into the tephra as it was deposited. The convoluted interface between the pre-Taupo sediment and the ignimbrite indicates that the emplacement disturbed and deformed the surface of the cave floor.

A small degree of sequence variation within species of moa was observed. It is impossible to attribute this variation either to post-mortem damage, polymerase errors, or genuine polymorphism, because sediments contain genetic material from an unknown number of individuals. This is one of the limitations of studying aDNA in sediments. A measure of damage could be obtained by Uracil-N-Glycosylase (UNG)-treating extracts before PCR amplification (Hofreiter et al. 2001) and comparing the cloned results with those obtained without UNG-treatment, but this is beyond the scope of this paper. Regardless of the cause of the sequence variation, the taxonomic affiliations of the sequences were unambiguous.

5.7.2 Plant DNA

Amplifications, cloning, and sequencing of plant DNA demonstrated the presence in the sediments of DNA from a diverse range of plant families (Table 5.3), all of which are still present in or near the study area. Although the plants could only be identified to family level, with species-specific primers it should be possible to test for the presence of particular species in temperate sediments.

5.7.3 Sheep DNA

The apparent chronological fidelity of the moa genetic profile contrasts with that of the sheep. Sheep DNA was successfully amplified from European and post-Taupo layers in both sites, and also from the Taupo ignimbrite in Hukanui Pool at 350 mm below the base of the European layer. These data clearly show that sheep DNA has migrated down through the sediments and the genetic chronology has been, at least in part, disrupted. One possible reason for the difference in apparent mobility between the DNA of moa and sheep is that birds, unlike sheep, do not generate large volumes of DNA-containing urine (Valiere and Taberlet 2000). It is likely that the copious amounts of sheep urine deposited on the surface carried sheep DNA downwards, rather than sediment particles themselves moving, for which there is no corroborating geological evidence. Quantitative PCR of the sheep DNA in the sediments showed that there was approximately eight times (three cycle shifts) more sheep DNA in the European Layer than in the underlying Taupo ignimbrite in Hukanui Pool, which supports the hypothesis of migration by transport in urine from above (Table 5.4).

Table 5.3 Distribution of plant families in Hukanui Pool and Hukanui #7a

Numbers in brackets represent the highest posterior probabilities over 70% obtained for that family. [a]*rbc*L H1af, *rbc*L H2ar, [b]Upper Carex 18-mer, Lower Carex 18-mer, [c]*trn*L*g*, *trn*L*gh*

	Hukanui Pool	
Strata	**Sample No.**	
F European	HPLF-70	Araliaceae(92%)[a],(100%)[b], Asteraceae(100%)[b], Fabaceae(99%)[a], Griseliniaceae(75%) [b], Plantaginaceae(100%)[b], Poaceae(100%)[b], Rosaceae(100%)[b]
E Post-Taupo	HPLE-150	Poaceae(100%)[a], Polygonaceae(96%)[a]
D Taupo	HPLD-90	Asteraceae(99%)[a], Coriariaceae(92%)[a]
C Pre-Taupo	HPLC-10	Onagraceae(94%)[b], Plantaginaceae(100%)[b], Polygonaceae(99%)[a], Ruscaceae(99%)[b]
B Waimihia	HPLB-10	Onagraceae(71%)[a]
A Pre-Waimihia		

	Hukanui #7a	
Strata	**Sample No.**	
F European/mixed	H#7a Layer F-3	Poaceae(100%)[c], Asteraceae(100%) [c], Lamiaceae(74%) [c], Onagraceae(100%)[c], Fabaceae(100%)[c], Juncaceae(100%)[c]
E Post-Taupo	H#7a Layer E-10 H#7a Layer E-80	Asteraceae(100%)[c] Onagraceae(100%)[c], Polygonaceae(90%)[b]
D Taupo	H#7a Layer D-100	
C Pre-Taupo	H#7a Layer C-160	Ranunculaceae(100%)[c]
B Waimihia	Not present	
A Pre-Waimihia	Not present	

Table 5.4 Results of qPCR sheep primers on Hukanui Pool sediments. Quantity of PCR amplifiable template within the extracts is inversely related to cycle threshold (CT) value, with a lower CT value indicating a greater number of initial template molecules. Under an assumption of perfect amplification kinetics, during the early phases of exponential growth in template number (*i.e.*, before limiting effects of resource abundance), the total quantity of PCR can be expected to approximately double per cycle. Therefore a difference of one cycle in the CT value of any two extracts within a comparable qPCR reaction can be roughly interpreted as an initial two-fold difference in template molecules. As such the extracts from level F contain approximately eight times more DNA than level D (approximate CT difference of three).

Strata	**Sample No.**	**Extraction No.**	**CT Value**	**Comments**
European	Hukanui Pool Layer F-70	e32	27.78	Correct Tm
Post-Taupo	Hukanui Pool Layer E-150	e43, e44, e42	n/a	Extract inhibition precluded any DNA amplification
Taupo	Hukanui Pool Layer D-20	e39	30.40	Correct Tm
Taupo	Hukanui Pool Layer D-20	e38	31.99	Correct Tm
Pre-Taupo	Hukanui Pool Layer C-10	e23, e24	n/a	No amplification of target DNA, as identified by anomalous Tm
Waimihia	Hukanui Pool Layer B-10	e36, e36	n/a	No amplification of target DNA, as identified by anomalous Tm
Pre-Waimihia	Hukanui Pool Layer A-10	e34, e35	n/a	No amplification of target DNA, as identified by anomalous Tm

Studies of modern sedimentary bacterial DNA show that a major part of the nucleic acids released from their cellular matrices after cell death quickly binds to the surfaces of quartz, feldspar, clay, humic acids, and other soil components, which reduces leaching and degradation of DNA by nucleases (Lorenz and Wackernagel 1994). Thus, it could be that the mineral surfaces that bind free DNA had been saturated, and an excess of DNA would be able to pass. It is also possible that a relatively large volume of urine would simply carry its DNA load through the sediment until the flow reached its limit. Even if one or both mechanisms have functioned, DNA from non-permafrost sediments may still be useful in reconstructing past processes, particularly where similar levels of soil DNA saturation have not been reached. However, each site will need to be assessed individually and success may depend on the sediment composition, the biota occupying the site, and events in the site's history (such as flooding).

5.8 Conclusion

Our results suggest that the distribution of amplifiable DNA preserved directly in the non-cryogenic sediments of these sites corresponds to sequences in the faunal and floral histories of the sites and their local environments reconstructed from other information. However, the results also suggest that considerable caution must be exercised in interpreting DNA profiles from sediments both because of downward movement of DNA in some circumstances, and of other factors that can govern the presence of DNA from a locally-occurring organism in a site or stratigraphic layer. Further studies need to be conducted on these sites to investigate sample-to-sample variation of DNA within strata and on other localities and sediment types to

investigate whether DNA from upper layers regularly penetrates the layer(s) below the occurrence of a taxon and, if so, to what extent this problem exists. The evidence presented here for organisms that do not produce copious amounts of liquid urine, suggests that most, if not all, of the DNA from such taxa is stratigraphically localised in the sediments. If the methodological challenges can be overcome, aDNA from non-frozen sediments may prove to be a powerful molecular tool in the reconstruction of past environments

Chapter 6

The Provenancing of Animal Remains from Archaeological Sites using Ancient DNA

provenance ('prəvənəns). [a. F. *provenance* (prəvnãs) *Dict. Acad.* 1835, f. provenant, pres. pple. of *provenir* to come forth, arise, ad. L. *prōvenīre*: see PROVENE.] **a.** The fact of coming from some particular source or quarter; origin, derivation.

Oxford English Dictionary

6.1 Abstract

Genetic provenancing has significant implications not only for fully characterizing the occupation history of archaeological sites, but also to provide novel insights into the origins and spread of domesticated animals, as well as the provenance of economically important wild fauna. To illustrate this approach, we present examples using DNA extracted from bones of dugong (*Dugong dugon*) and pig (*Sus scrofa*) from an archaeological site at Fort Frederik Hendrik, Mauritius and pig bones from two sites on Lord Howe Island, Australia. The recent increase in both the abundance and taxonomic range of DNA sequence data in public repositories makes it possible to determine the provenance of faunal archaeological material by characterizing its DNA. Among the most commonly represented taxa are domesticated animals, for which extensive genetic characterization has revealed high levels of genetic diversity and strong phylogeographic clustering. As many of the animal remains found at Neolithic and more recent archaeological sites are likely to belong to domesticated animals, DNA extracted directly from this material can be used to determine the probable geographic origin of that specimen.

6.2 Introduction

Determining the provenance of archaeological finds is one of the corner-stones of archaeology, as it can reveal a wealth of information about economic activity, trade routes, resource exploitation, material life histories, and social and political systems. Whilst archaeological provenancing techniques for inorganic materials such as pottery, stone axes and metal tools are well developed, studies of organic remains have tended to be restricted to morphological analysis and/or characterization of their

inorganic constituents (stable isotopes and trace elements). However, in very recent years, it has become possible to extract and sequence DNA from subfossil remains of various organisms. Such ancient DNA (aDNA) techniques have been used to provenance museum specimens (Austin et al. 2002; Barnett et al. 2007), to identify the species from which archaeological animal bone was derived (Newman et al. 2002), to provenance ancient archaeological plant material (Deguilloux et al. 2002; Haneca et al. 2005), to interpret prehistoric human movements by proxy through the amplification of the DNA of domestic and commensal animals (Matisoo-Smith *et al.* 1998; Larson *et al.* 2007). It is therefore a logical step for aDNA of economically important archaeologically derived faunal remains to be used in provenancing.

Two criteria must be met for this to be practical. Firstly, a large mitochondrial reference database of a specific target species must be available. Mitochondrial DNA (mtDNA) is favoured over nuclear DNA for aDNA studies as many more copies are present within the cells, thereby increasing the probability of survival and recovery from bones and other organic remnants. In addition, the lack of recombination in mtDNA facilitates phylogenetic and demographic reconstruction using the recovered sequences. However, given the rapid rate of post-mortem DNA decay (Hoss et al. 1996), the difficulty of amplifying long tracts of aDNA and the relatively recent evolutionary history shared by common descendents of domesticated animals it is necessary to target a small yet divergent portion of the mitochondrial genome for analysis. To meet these criteria, the most common genetic target for aDNA provenancing is the hypervariable region within the mitochondrial control region: a segment of the mtDNA genome that does not code for any proteins, should not be

influenced significantly by natural selection, and consequently exhibits a rapid evolutionary rate compared to other regions in the mtDNA and nuclear genome.

Secondly, the genetic signature recovered from the sample must be characteristic of a restricted geographic region, as occurs when species are associated with a strong phylogeographic signal. Correlation between genetic signatures and geographic provenance has been demonstrated for pigs (Larson et al. 2005), cows (Troy et al. 2001), rabbits (Hardy et al. 1995), goats (Kahil Bar-Gal et al. 2002), sheep (Guo et al. 2005; Kijas et al. 2006), and to some extent for chickens (Liu et al. 2006). Additionally, while some phylogeographic signal has been identified in horses (Vila et al. 2001) and dogs (Savolainen et al. 2002) it is significantly less strong than that observed for the other taxa. In the case of the pig, which is the most thoroughly genetically-described species, a single nucleotide polymorphism is sufficient to place an individual within a specific and circumscribed region of the world.

Oceanic islands present an ideal opportunity to explore the potential of aDNA provenancing. Prior to human settlement, oceanic islands would have been devoid of domestic species, and as with human colonisation, it is not always possible to determine the origin of the introduced animals from the material remains alone.

6.3 Sites Descriptions and Locations

6.3.1 Mauritius

One such oceanic island is Mauritius, which is located in the Southwest Indian Ocean approximately 900km east of Madagascar. Before the arrival of humans, Mauritius had no domesticated animals and it may have been the Portuguese, who first visited Mauritius in 1507, that introduced them. Due to its key position along the trade route, frequent movement back and forth between Mauritius and other islands resulted in the transfer not only of people and goods but also of various animal species. Giant tortoises (*Dipsochelys elephantina* and *Asterochelys radiata*) were introduced to Mauritius in the 1830s (Stoddardt et al. 1979), deer (*Cervus timorensis russa*) in 1637 (Grimaud et al. 2006), goats in the 17th Century (Cheke and Hume 2008) and Crab Eating Macaques (*Macaca fascicularis*) from the Indonesian Archipelago, in 1606 (Cheke 1987; Cheke and Hume 2008).

By contrast, dugongs were one of the species endemic to Mauritius and were last mentioned by inhabitants of the island in 1799 after which they became locally extinct (Staub 1993; Cheke and Hume 2008) through overhunting. During Dutch times the dugong was praised for its "delicious meat, one of the best that could be found on land or sea" and in the same document from 1674 the Dutch commander Hubert Hugo added that the dugong could be salted or smoked (Staub 1993; Cheke and Hume 2008).

Although the Dutch mention the damage and abundance of feral pigs, goats and cattle they failed to mention anything about the endemic fauna even after a decade of colonisation of Mauritius. However, the fort's faunal remains provides insights into

the diet of the colonists, as well as into their impact on the local species. Preliminary results indicate that the main meat source was deer, followed by goat or sheep and pig, and a small amount of cattle. While it was supposed that the colonists would have incorporated the local fauna into their diet, such as the now-extinct dodo (*Raphus cucullatus*) and giant tortoise (*Cylindraspis* sp.), no endemic terrestrial species have been identified in the faunal record at the fort with the exception of a tibiotarsus of the extant Pink Pigeon (*Nesoenas mayeri*) (Hume unpub.) found within the deepest layers of the settlement. This is presumably due to the rapid extermination of endemics within habited areas and the increase in readily available feral stock.

In 1638, the Vereenigde Oost-Indische Compagnie (VOC *or* Dutch East India Company) began construction of Fort Frederik Hendrik, a four-sided star fort overlooking the southeastern harbour, the present day Vieux Grand Port. One meter deep excavation trenches at Fort Frederik Hendrik, were dug within colluvial sediments which are up to 1.5m thick and overlie bedrock, with the groundwater table well below the soil stratum. These sediments are mixed with the material residues of the colonists, such as fragments of ceramics and building materials as well as anthropogenically added bone layers and dispersed shell fragments which has the effect of raising the pH to values higher than 7 and thus preventing dissolution of the bones and collagen.

From the Fort Frederik Hendrik site we focused on skeletal remains that had previously been morphologically identified as either pig or dugong, as a mitochondrial DNA database for comparative purposes already existed for both of these species, and both demonstrate strong phylogeographic clustering. Furthermore,

for these two species, provenance is indeterminable from morphological evidence alone. We took samples of sediment from areas immediately adjacent to the sampled bones.

6.3.2 Lord Howe Island

We also obtained DNA sequence data from faunal material excavated at two sites on Lord Howe Island, a small remnant volcanic island located between the Australian east coast and New Zealand in the Tasman Sea (Figure 6.1). The island was uninhabited prior to its discovery in 1788 (Rabone 1940; Anderson 2003) but rapidly became a resupplying destination for passing whaling, sealing, trade and colonial ships. These ships included locally owned vessels which operated around the Tasman Sea, those frequently trading between the colonies and Asia, and long-distance traders and whalers from Britain and the Americas. The ship traffic to the uninhabited island during the first 40 years following its discovery resulted in the release of pigs (*Sus scrofa*) and goats (*Capra hircus*). As a result of these introductions some native species such as the Lord Howe Swamphen or White Gallinule (*Porphyrio albus*), the White-throated Pigeon (*Columba vitiensis godmanae*), Red-fronted Parakeet (*Cyanoramphus novaezelandiae subflavescens*) and the Tasman Booby (*Sula tasmani*), had become extinct or highly endangered by the time permanent settlement began in 1834 (Hutton 1990).

Originally settled by a small group comprised of three British ex-whalers, their Maori wives and children, Lord Howe was subsequently colonised with the aim of formalising its role as a place of refreshment. Employed by a local Australian whaling firm, the settlers used the island as a supply depot by establishing a small scale

agricultural industry (Nichols 2006). Staple crops such as potatoes, onions, cabbages and kumara were introduced and grown expressly for the ship trade, while local fish, mutton bird and feral game were offered for trade as well as to provide a significant portion of the local subsistence. Pigs and goats already on the island were incorporated into this enterprise, but whether the initial settlers brought further pig and goat stock along with their fowl (*Gallus* and *Anser*) and dogs (*Canis*) is unclear (Rabone 1940). The supply trade continued on the island until the increased availability of petroleum products in the mid 1850s precipitated a global slump in the whaling industry (Mawer 1999). By the 1870s the ship trade was almost finished and the Lord Howe economy was forced to move away from the supply trade in favour of a couple of (failed) agricultural exports, and later the successful export of an endemic palm tree as an indoor ornamental plant for the European market (Rabone 1940; Nicholls 1952).

Archaeological investigations of the settlement landscape of Lord Howe focused on midden materials from two domestic occupation sites, OSBF (-31° 30' 51", 159° 3' 51") and PJL (-31° 32' 40", 159° 5' 15") (Figure 6.1). OSBF is the earliest site on the island, dating from around 1834 to about 1845-50 and was most probably built and occupied by the first settlers who came directly from New Zealand shore whaling stations. These first settlers may have brought pigs with them from New Zealand, or could equally possibly have been supplied with animals from the Australian colonies. There are also indications that by this time there was already a significant feral pig population established on the island, which was descended from animals released on at least two occasions by British, American or local Colonial ships that could have brought pigs from anywhere in the Pacific/Asia or Australian/New Zealand colonies

to Lord Howe Island. Excavations of the earliest occupation layers are poor in post-cranial pig and goat remains, indicating that perhaps the bulk of the meat producing portions of these animals were exported from the island as articles of trade.

The PJL house site dates from the 1860s to about 1920 and contained an increased variety and frequency of post-cranial pig, goat and other domesticate skeletal elements, indicating that more meatier portions were remaining on the island more often, rather than being traded off as smoked and salted meats. By the time of the establishment of this site, shipping was on the decline, although there remained ample opportunity for supplementary domestic stock (particularly pigs) to be brought to the island; either as specific imports from the colonies, China or Britain, or as opportunistic trade items with passing ships who had a surplus of livestock but not vegetables (Rabone 1940; Nicholls 1952). As shipping decreased further, opportunities for new domestic stock arrivals became fewer and the general incidence of ships from outside the local region became much less.

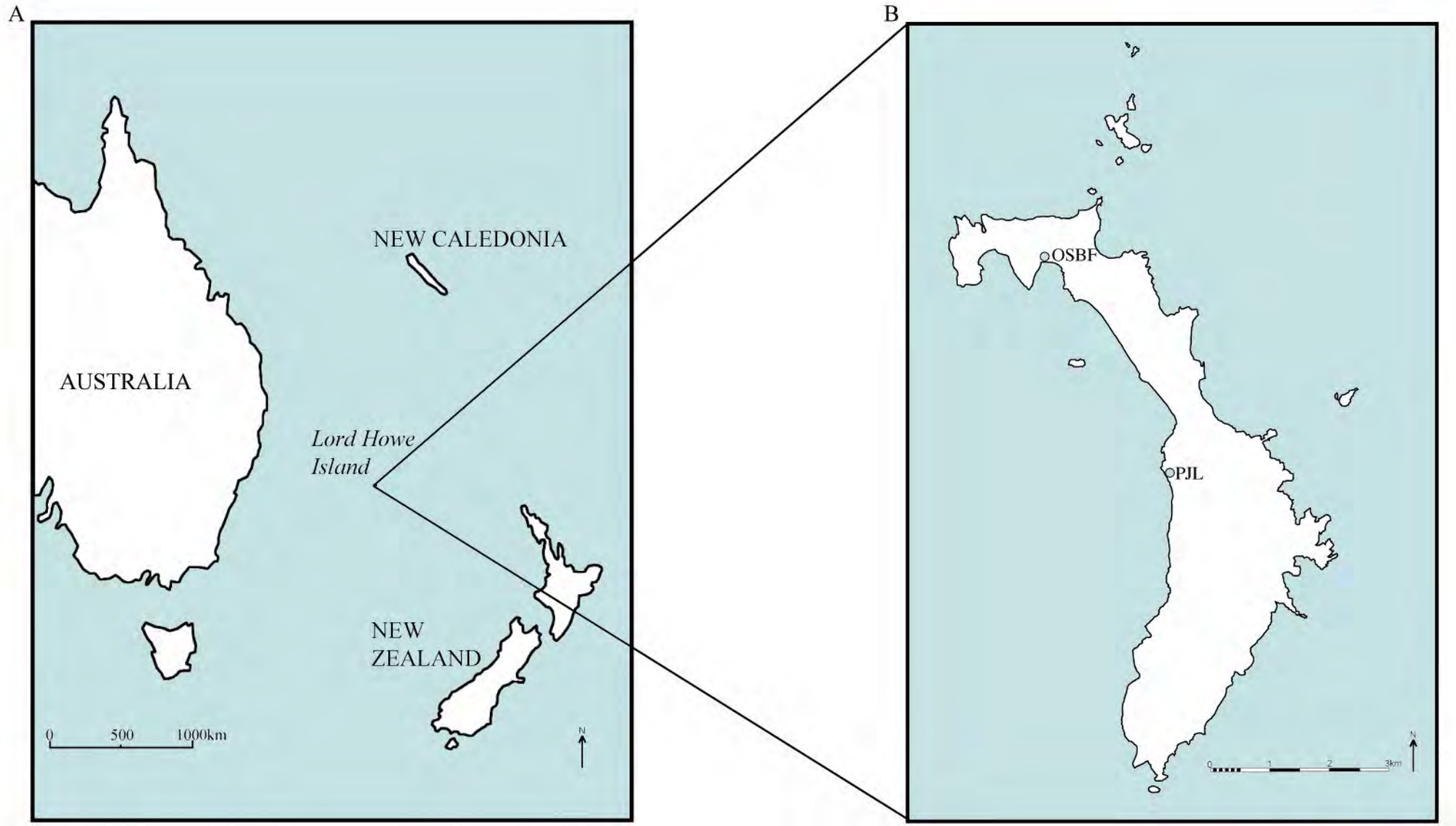


Figure 6.1 Two maps depicting (A) the position of Lord Howe Island within the Tasman Sea, and (B) the location of the OSBF and PJL sites on the island.

6.4 Samples

6.4.1 Mauritius Samples

One pig and two dugong bones were collected from deposits on the western slope of the plateau on which Fort Frederik Hendrik once stood, containing building debris and domestic refuse, ranging from Dutch clay tobacco pipes to Japanese porcelain (excavation trench 26, feature 3547). These artefacts date from c. 1675-1710, the last decades of Dutch presence on the island.

Ancient DNA of animal origin has been successfully extracted from permafrost as well as temperate sediments (Willerslev et al. 2003; Haile et al. 2007). However, to date there have been no published reports of sedimentary aDNA (sedaDNA) being successfully extracted from tropical sediments, and to test if this substrate could be a possible source of DNA approximately 10g of sediment samples were taken directly adjacent to the collected bones.

6.4.2 Lord Howe Island Samples

Four pig teeth from kitchen midden deposits at OSBF and four pig teeth from domestic underfloor deposits at PJL were collected from occupation levels dating from the 1830s to the 1840s (OSBF) and the 1860s to 1920s (PJL). Both midden and underfloor deposits also contained a wide selection of domestic refuse, including a large variety of faunal remains from domestic and native species, common 19th century British household ceramics and glassware, Scottish clay tobacco pipes, and common beads, buttons and ammunition. From their history and the known demographics of occupants, these two sites constitute a representative sample of most

households on the island spanning from first settlement to the beginning of mature settlement at the beginning of the 20th century.

6.5 Methods

Appropriate sterile techniques were followed to avoid potential contamination of the samples either with DNA from other samples or from modern humans (Cooper and Poinar 2000). DNA was extracted, amplified and sequenced using standard aDNA techniques that have been optimized for bone material (Shapiro et al. 2004), in a dedicated aDNA facility at Oxford University and following established aDNA protocols (Cooper and Poinar 2000).

The sediment samples were extracted using the following protocol: each sample was placed in PowerMaxTM Soil PowerBead tubes (Cambio UK Ltd), and dissolved in 24 ml lysis buffer (Bulat et al. 2000). The tubes were then agitated vigorously for 1 min and rotated overnight at 65°C. Following extraction, the DNA was purified according to the PowerMaxTM Soil DNA Isolation Kit protocol, before being concentrated with Amicon Ultra 30 kDa spin columns.

To assess whether DNA survived in the samples, 2µl of DNA extract was amplified using PCR with 50 cycles of 1.5 min initial denaturation at 97 °C, 45 s at 94 °C, 45 s at 56–60 °C, 1.5 min at 68 °C and a final cycle of 10 min at 68 °C. Primers used are listed in Table 6.1. Extraction blanks and PCR negative controls were used throughout the experimental procedures. No negative control contained any amplification product despite the large number of cycles used, and the potential problem of contamination from modern domesticated animals (Leonard et al. 2007).

Table 6.1 Showing primers used to amplify pig and dugong control region mtDNA.

Primers	**Target Group**	**Sequence**	**Reference**	**Annealing Temp. (°C)**	**Product length (bp)**
Anc f1	Pig	5′-CTTTAAAACAAAAAAACCCATAAAAA	(Larson *et al.* 2007)	56	125
Anc r1	Pig	5′-TTAATGCACGACGTACATAGG	(Larson *et al.* 2007)	56	
Anc f2	Pig	5′-CGTACATAGCACATATCATGTCA	(Larson *et al.* 2007)	56	121
Anc r2	Pig	5′-AGAAGAGGGATCCCTGCCAA	(Larson *et al.* 2007)	56	
PNG f	Pig	5′-TAATCAGCCCATGCTCACAC	(Larson *et al.* 2007)	56	104
PNG r	Pig	5′-ACTGTGTTAGGGCCTTTGAC	(Larson *et al.* 2007)	56	
Dugong CR F2	Dugong	5′-TATTACAACGGTCTTGTAAACC	*This study*	56	270
Dugong CR R1	Dugong	5′-GGCTTTAGAGCATGAATCCTA	*This study*	56	
Dugong CR F1	Dugong	5′-CATCCTATGTATAATCGTGCAT	*This study*	56	266
Dugong CR R2	Dugong	5′-GGGTTGCTGGTTTCTCGAA	*This study*	56	
Dugong CR F3	Dugong	5′-ATAGTCCTTGTCAACACGCTT	*This study*	56	230
Dugong CR R3	Dugong	5′-GTCATAAGTCCATCGAGATGTC	*This study*	56	

PCR products were cleaned using a QIAquick PCR Purification Kit (QIAGENE). Dugong amplification products were cloned using the TOPO TA cloning kit from Invitrogen, and eight colonies picked. Amplification products were cleaned by ethanol precipitation, and the pellet was then transferred to the DNA sequencing facility (Zoology Department, University of Oxford) where it was sequenced on an ABI 3700 according to manufacturer's instructions.

DNA extracts from pig bones were first subjected to PCR amplification using the Anc f1 / Anc r1 primer pair. These two primers bracket a region of the mitochondrial DNA control region which possesses diagnostic single nucleotide polymorphisms (SNPs) that discriminate not only between European and Asian haplotypes, but also between individual European haplotypes. In order to differentiate between the samples which possessed an Asian signature using the first primer pair, two additional PCRs were carried out using two sets of primer pairs: Anc f2 / Anc r2 and PNG f / PNG r. Additional details regarding the differences between all the European and Asian haplotypes are discussed in Larson et al 2007a and Larson et al 2007b respectively (Larson *et al.* 2007; Larson *et al.* 2007).

Dugong sequences were aligned with a dugong (Blair et al. in prep.) reference sequence using ClustalW in BioEdit (Hall 1999). A maximum likelihood tree (Figure 6.2) was constructed using PAUP (Swofford 1993) with the HKY+G model of nucleotide substitution. The initial value for the rate heterogeneity parameter as well as the tr/tv ratio was estimated from a NJ tree and fixed. The ML tree was then estimated using NNI branch swapping. Model parameters were then re-estimated and fixed, followed by an iteration of TBR branch swapping. Support for the tree structure

was evaluated through 1000 full heuristic bootstrap replicates which were then performed, with model parameters set to the estimated values, starting trees generated by NJ, and NNI branch swapping. Additional support for the evolutionary relationships among the sequences was evaluated and Bayesian posterior probabilities estimated using the program MrBayes (Hulsenbeck and Ronquist 2001). Two chains were run for 10 million generations each, again using the HKY+G model of nucleotide substitution. Nodes with Bayesian posterior probabilities >80% and ML bootstrap consensus scores >70% (bold) are indicated in the tree.

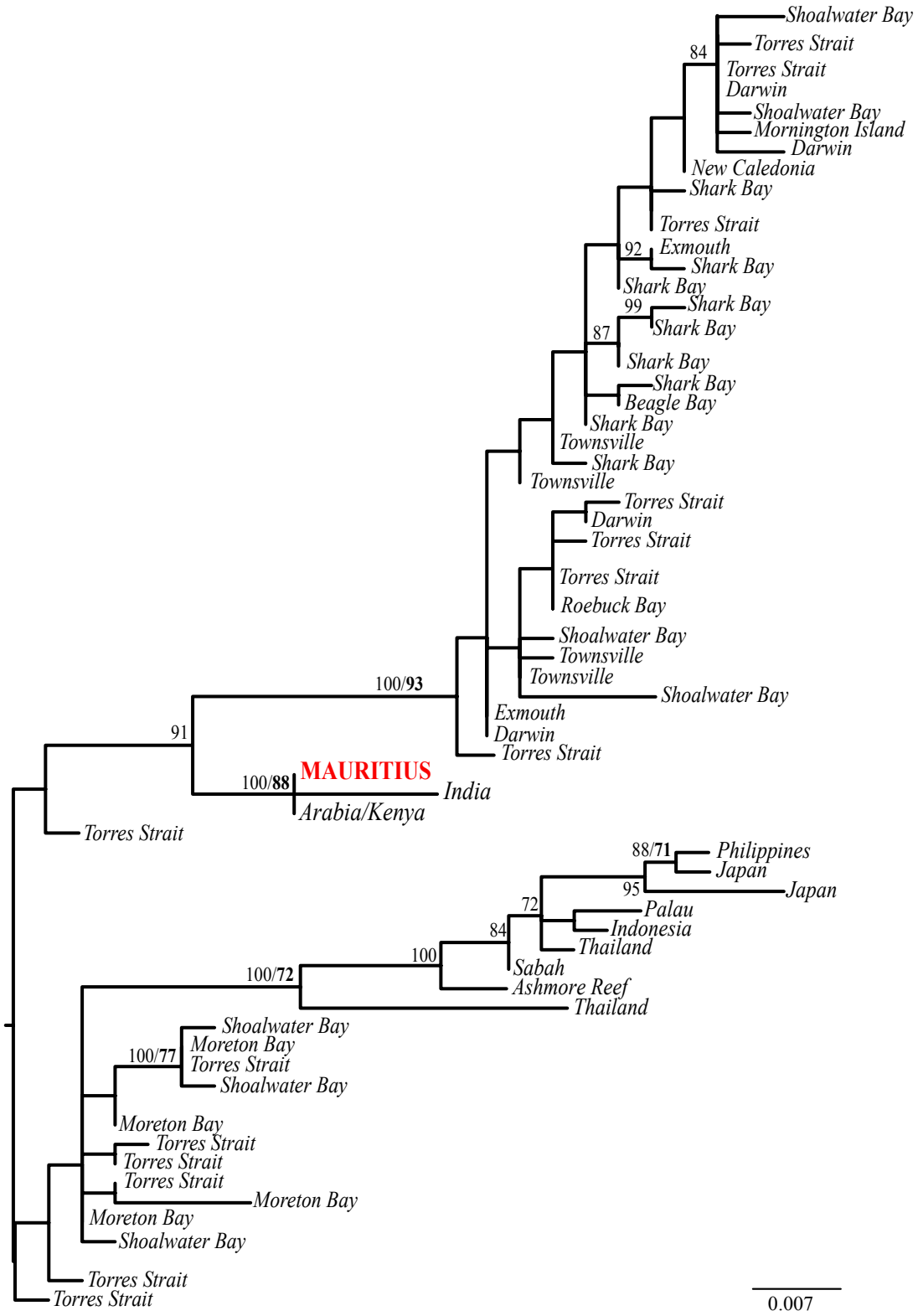


Figure 6.2. A midpoint rooted maximum likelihood tree of dugong control region (partial) sequences. Tip labels refer to the locations from where the samples were sourced. Node lables in normal type are boostrap values, and in bold type are posteria probablities.

6.6 Results

6.6.1 Mauritius Dugong

Two dugong rib bones lying next to one another were retrieved from the Fort Frederik Hendrik deposits. The sequences derived from these bones were identical (GenBank Accesion Nos. EU826002-3) (Table 6.2), and the bones may have belonged to the same individual. Comparison to other published dugong sequences (GenBank accession Nos. EU835761-EU835816) (Blair et al. in prep.) revealed that the Fort Frederik Hendrik sequences are phylogenetically distinct from Australian and East Asian sequences, but are identical to a short (194 bp) sequence from two Arabian dugongs and one Kenyan animal, and also cluster with that of an individual from India (Figure 6.2).

6.6.2 Mauritius Pigs

The sequence derived from the Mauritian pig bone (Table 6.2), GL1077, retrieved from Fort Frederik Hendrik was identical to haplotype 'C' (Larson *et al.* 2007), the second most common haplotype found amongst European domestic pigs, and a haplotype not found on any other continent (Table 6.3).

6.6.3 Lord Howe Island Pigs

All four pig teeth (GL1093-1096) from the Lord Howe Island site OSBF yielded sequences with the most common domestic European haplotype, 'A' (Table 6.2, Table 6.3)(Larson *et al.* 2007).

From the PJL site, all four pig teeth (GL1097-1100) possessed East Asian haplotypes as determined by the first primer pair. Sequences derived from the 2nd and 3rd primer pairs, however, demonstrated that each sample possessed a different haplotype (Table

6.2, Table 6.3). The specific results are as follows. For haplotype designations see Larson et al. 2007b (Larson *et al.* 2007):

GL1097 was identical to four of the most common East Asian domestic pig haplotypes (GX, L, DW, EH) found in specimens from China, Australia, Guam, and Hawaii (Table 6.2, Table 6.3);

GL1098 possessed a haplotype which was novel, but still distinctively East Asian (Table 6.2, Table 6.3);

GL1099 possessed a diagnostic SNP in the third primer pair indicating that it was identical to the most common Pacific Clade haplotypes (BX) found in modern and ancient material derived from New Guinea and throughout Near and Remote Oceania, and associated with the initial movement of humans into the Pacific (Table 6.2, Table 6.3);

and GL1100 possessed a common East Asian haplotype and was identical to the D and AE haplotypes found in China and in Guam (Table 6.2, Table 6.3).

The presence of the Pacific Clade haplotype on Lord Howe is interesting since no domestic pigs on either New Zealand nor Australia have yet to be shown to possess a Pacific signature, despite the ubiquity of Pacific Clade pigs in New Guinea and in the islands within the Near and Remote Pacific.

6.6.4 Mauritius Sediments

No pig or dugong DNA were successfully amplified from the sediments sourced immediately adjacent to the bone samples.

6.7 Discussion

6.7.1 Mauritius

It has been noted that DNA from bones is best preserved under dry, cold conditions, and biochemical studies suggest that DNA is unlikely to survive longer than 100 000 years (Wayne et al. 1999). However, this is a crude estimate, and rather than absolute age of a sample, DNA survival is likely to be a function of time and temperature, and it is therefore more reasonable to evaluate DNA persistence in terms of thermal age (Smith et al. 2003). Similarly, it has been shown that DNA survival can also be related to latitude (Karanth et al. 2005), presumably correlated with thermal age.

An increase of temperature of 20 °C is estimated to result in a 10-25 fold increase in DNA decay by hydrolysis and oxidation (Hoss et al. 1996). As the site of Fort Frederik Hendrik is situated in the tropics at 20° 22' S; 57° 43' 5 E, with a high annual mean temperature of 23 °C (Padya 1989), and the dugong and pig bones were not excavated from a cave site, it is initially surprising that aDNA was recovered at all. Mitigating the effects of temperature and latitude, is the relatively young age of the samples (300 years old). In addition anthropogenically added Calcium in the form of shell and bone material as observed in the Fort Frederik Hendrik trenches raised its pH to values higher than 7 and prevented dissolution of the bones.

By contrast, animal bones recovered from a nearby waterlogged marsh, the Mare aux Songes, and dated to approximately 4000 years BP did not yield any aDNA (Rijsdijk et al. (in prep.)). Based on published reports of successful aDNA studies, 2000 years appears to be the upper limit for aDNA survival in the tropics (Karanth and Yoder in

press). This is not necessarily a concern for tropical archaeological studies, as the archaeological period of interest in most tropical oceanic islands fall within this window of aDNA survivability.

Whereas the skeletal remains yielded aDNA, the sediments did not. Following release of DNA from an organism into the soil, it is vulnerable to oxidative and hydrolytic damage and is readily attacked by microbial DNases and nucleases (Crecchio and Stotzky 1998), whose activity is positively correlated with soil temperature (Gulden et al. 2005). However, the binding of extracellular DNA to soil particles, such as clay minerals, humic acids, and silica particles, can confer protection against these enzymes (Lorenz and Wackernagel 1994). Thus the survival of DNA in sediments is a function of the characteristics of the soil, such as the relative proportions of different clay minerals, cation concentration, pH and soil temperature.

The sediments from which the bone samples were extracted are clayey colluvial soils with a pH above 7. The presence of free water in the burial environment is essential to the function of bacterial DNases, believed to be the primary cause of extracellular DNA degradation (Blum *et al.* 1997). Although sedaDNA was recovered from New Zealand rock-shelter sediments that ranged in age from 3000 BP to present, there was no free water in these sites, and the annual mean temperature is less than in Mauritius (Haile et al. 2007).

It is of course equally possible that there was never any pig or dugong DNA in the sediments, as the bones might have been deposited after they had been de-fleshed and that no DNA leaching from the bones has occurred. It is also likely that pigs were not

penned or butchered in that part of the fort and so could not have contributed DNA to the sediments through their faeces, urine, hairs, skin cells etc (Willerslev et al. 2003).

As well as introducing animals to Mauritius, Dutch colonists also exported animals as food for sailors. These included live tortoises, and dodo, and even salted dodo that they could not immediately eat, to use as a shipboard provision (Quammen 1996). Within this context of inter-island movement of animal species, the origin of the dugong and pig remains is not immediately apparent.

Since the dugong sequences are identical to those of animals from Arabia and Kenya, it is conceivable that the Mauritian remains could have originated there, transported as preserved meat provisions by Dutch sailors from other parts of their trading empire, although unlikely given the abundance of local dugong in the surrounding lagoons. Another possibility is that sailors carried with them dugong bones to carve into scrimshaw. However, this similarity with Arabian and Kenyan samples is based on a relatively short sequence, and it is possible that local dugong populations also had an identical genetic signature in this part of the mtDNA genome. Since dugong were once common throughout the Mascarenes (Cheke 1987; Cheke and Hume 2008), it is more likely that they were exploited locally. Nevertheless this finding has implications for any future conservation or possible reintroduction program.

As for the Mauritius pigs, it had been presumed that the bones recovered from the site of Fort Frederik Hendrik came from locally butchered animals that had been introduced onto the island from Europe and which had quickly become feral.

Alternatively, it is possible that the Dutch imported pigs from their colonies in India and South-East Asia.

However, the results from the pig bone shows that at least some pigs were brought from Europe, via the Cape where a large Dutch presence had been in place since 1652 (Moree 1998). The 'C' haplotype found is diagnostic for European pigs and wild boar, and quite clearly distinct from those of East Asia. Although the mtDNA results reflect the maternal lineage, the age of the remains predates the initiation of efforts to create improved breeds by mating imported East Asian domestics with European breeds, and so the maternal lineage in this case is sufficient to claim that the pig is from Europe. On other islands, this might not be the case, therefore this result demonstrates the effectiveness of this methodology in determining provenance when a reliable chronology is also known.

6.7.2 Lord Howe Island

The finding that two sites lying just 8 km apart contain pigs with exclusively Asian or European haplotypes is surprising, especially since the material culture of both settlements is essentially identical. Of course, with such a small sample size, the difference could merely be the result of chance, of founder effect, rather than a conscious choice made by the settlers.

An alternative explanation is that the European haplotypes, which are found only at the earlier OSBF site, could have originated from ships that had recently resupplied in the colonies. This could apply regardless of whether they are from the feral stock first released on the island, or later introductions by the first colonists or early ship trade. In contrast, by the time of the establishment of PJL, there would have been sufficient

time for any later stock introductions from other sources to have become established on the island; either as new additions to feral populations, or as husbanded animals. Further, although shipping was on the decline by this point, what shipping there was typically came from all over the Pacific (Rabone 1940; Nicholls 1952), providing ample opportunity for the introduction of Asian pigs. The far-reaching nature of Lord Howe's maritime network and consequent opportunities for stock recruitment are exemplified by the stories of two of settlers who arrived during this time, including the man who settled PJL, Perry Johnson. Perry Johnson was an emancipated American slave who had come to the island as a crewman on an American whaler, the Will o' the Wisp which would have had an extended and convoluted voyage throughout the Pacific as was the norm for American whaling vessels at this time (Rabone 1940; Nicholls 1952; Mawer 1999). Similarly, Nathan Thompson of the American whaler Belle disembarked and settled at Lord Howe with three Gilbert Island (Kiribati) women who had been picked up during his ship's previous voyaging around the Pacific, which also included recorded stops in Vanuatu and New Caledonia (Rabone 1940; Nicholls 1952; Nichols 2006).

6.8 Conclusions

The examples taken from three sites in two oceanic islands demonstrate that it is possible for the first time to determine the provenance of species that have a well-characterised phylogeographic signal in conjunction with a mtDNA reference database. This materially increases the amount of archaeological information that can now be obtained from skeletal material, and is perhaps of particular interest to those studying recent human settlement of oceanic islands. Furthermore, when combined with genetic and morphological sex identification, ancient DNA provenancing may

also reveal differences in animal husbandry and the bias of movement of animals of a particular sex.

Table 6.2 Showing raw sequence data of Dugong and Pig Control Region. ~ indicates missing bases.

Sample	**Sequence**
Mauritius Dugong	AATTCTACTTAAACTACTCCCTGCGCGCGCTATGTACTTCGTGCATTATGTGCTCTTCCCCATAACAGTACTATATATGTTTTATC TTACATACACCATCCTATGTATAATCGTGCATTACACTACTTACCCCATGCATATAAGCCAGTACAGTAGGATTCATGCTCTAAA GCCTAAACAACTAACTTCCATTATACAGTCTCCACACCATGGATATTGTCCAGTCCATGTACCTCTTGATCTTGCATAGTACATTC AACCCTTTGTCGTACATAGCACATCTCTGAGATAGTTCTTGTCAACACGCTTATCACCTCCAATGGACAGTCCTTGACTACCAAG CTTCGAGAAACCAGCAACCCGCTCCGATTATGCTTCTCTTCTCGCTCCGGGCCCATAACTTGTGGGGG
Mauritius Pig ‘C’	TTGCGCACAAACATACAAATATGTGACCCCAAAAATTTAACCATTGAAAACCAAAAAATCTAATATACTATAAC
Lord Howe Island Pig ‘A’	TTGCGCACAAACATACAAATATGTGACCCCAAAAATTTTACCATTGAAAACCAAAAAATCTAATATACTATAAC
Lord Howe Island Pig ‘GX, L, DW, EH’	TTGCGCACAAACATACAAATATGTGACCCCAAAAATTTTACCATTGAAAACCAAAAAATCTAATATACTATAAC
Lord Howe Island Pig ‘East Asia’	TTGCGCACAAACATACAAATATGCGACCCCAAAAATTTAACCATTAAAAACAAAAAATTTAATATATTATAGC~AATAATTCCA GTCAACATGCGTATCACCACCATTAGATCACGAGCTTAATTACCATGCCGCGTGAAACCAGCAACCCGC~ATAACTGAGGTTTC ATACATTTGGTATTTTTTAATTTTTGGGGATGCTTGGACTCAGCCATGGCC
Lord Howe Island Pig ‘BX’	TTGCGCACAAACATACAAATATGCGACCCCAAAAATTTAACCATTAAAAACAAAAAATTTAATATATTATAGC~AATAACTCCA GTCAACATGCGTATCACCGCCACTAGATCACGAGCTTAATCACCATGCCGCGTGAAACCAGCAACCCGC~ATAACTGAGGTTTC ATACATTTGGTATTTTTTAATTTTTGGGGATGCTTGGACTCAGCTATGGCC
Lord Howe Island Pig ‘D, AE’	TTGCGCACAAACATACAAATATGCGACCCCAAAAATTTAACCATTAAAAACAAAAAATTTAATATATTATAGC~AATAACTCCA GTCAACATGCGTATCACCACCATTAGATCACGAGCTTAATTACCATGCCGCGTGAAACCAGCAACCCGC~ATAACTGAGGTTTC ATACATTTGGTATTTTTTAATTTTTGGGGATGCTTGGACTCAGCCATGGCC

Table 6.3. Showing haplotype designations of pig samples from Fort Frederik Hendrik and Lord Howe Island

Site	Sample Code	Elements	Notes	Extraction No.	Haplotype Designation	Geographic Region of Origin	GenBank Accession No.
Fort Frederik Hendrik, Mauritius	FH06-01	Bone	Trench 26	GL1077	C	European	EU826004
OSBF, Lord Howe Island	J5/I5/2-3/011	M1/M2	Maxillary - worn	GL1093	A	European	EU826005
OSBF, Lord Howe Island	J5/2-3/067	Premolar	Maxillary - worn	GL1094	A	European	EU826006
OSBF, Lord Howe Island	J5/2-3/083	Premolar	Mandibular premolar P4?	GL1095	A	European	EU826007
OSBF, Lord Howe Island	J5/2-2/132	Canine	Female	GL1096	A	European	EU826008
PJL, Lord Howe Island	2/3/012	Incisor	Mandibular	GL1097	GX, L, DW, EH	Asia (China, Australia, Guam, Hawaii)	EU826009
PJL, Lord Howe Island	PJ2/1/5/019	Premolar	Mandibular	GL1098		East Asia	EU826010
PJL, Lord Howe Island	A/3/035	Mandible	inc dP4, M1	GL1099	BX	Asia (Pacific)	EU826011
PJL, Lord Howe Island	A2/3/075	Premolar	Maxillary	GL1100	D, AE	East Asia (China, Guam)	EU826012

Chapter 7

Ancient DNA from Madagascan Elephant Bird Eggshells

All feathered things yet ever knowne to men,
From the huge Rucke, unto the little Wren;
From Forrest, Fields, from Rivers and from Pons,
All that have webs, or cloven-footed ones;
To the Grand Arke, together friendly came,
Whose severall species were too long to name.

Michael Drayton, 1604

7.1 Abstract

Avian eggshell is a commonly found artefact in archaeological and palaeontological sites, owing to its very recognisable structure and its resistance to decay. Because of the robustness of ratite eggshell, this material in particular has received much attention and is a rich source of information. Here we use preliminary results from ancient DNA from eggshell to shed light on the phylogenetic relationship of elephant birds and demonstrate the potential of this resource to the fields of phylogeogeography, population structure and environmental change.

7.2 Introduction

In archaeological sites the relative abundance of eggshell compared to other animal remains has been used to indicate changes in resource use (Sampson 1994); morphological variation in eggshell has helped to reveal taxonomic affinities (Harrison and Msuya 2005) and provide relative dating (Stidham and Brophy 2007); amino acid and stable-isotope biochemistry to examine palaeoecology (Johnson et al. 1997; Clarke et al. 2006); amino acid recemization for dating (Miller et al. 1999); and radioactive isotopes for radiocarbon dating (Freundlich et al. 1989; Higham 1994; Vogel et al. 2001).

Previous studies have shown that it is possible to extract DNA from modern avian eggshell in order to gain insights into parentage (Strausberger and Ashley 2001), and many studies have demonstrated that ancient DNA (aDNA) can survive in various biological substrates, such as hair (Higuchi et al. 1988; Gilbert et al. 2004), bone (Bunce et al. 2003; Bunce et al. 2005), tissue (Higuchi et al. 1984), as well as inorganic materials like sediments (Willerslev *et al.* 2003; Willerslev *et al.* 2004; Haile *et al.* 2007) and glacial ice (Willerslev et al. 2007).

This paper focuses on ratite eggshell of elephant birds (Aepyornithiformes) from Madagascar. The largest species of elephant bird, *Aepyornis maximus*, was 2.7 m tall, and whilst not as tall as the largest of the extinct New Zealand moa (Wetmore 1967), it was more massive, weighing in excess of 400 kg (Amadon 1947). Like all ratites, elephant birds were flightless, with no keel, possessed vestigial wings, long legs, and a long neck, and were presumably vegetarian, at least when adult.

The most recent revision of the order, based on skeletal morphology, has four different species of *Aepyornis*: *A. maximus, A. medius, A. hildebrandti,* and *A. gracilis* (Brodkorb 1963). The genus *Mullerornis* is also thought to contain four species: *M. betsilei, M. grandis, M. agilis,* and *M. rudis* (Hawkins and Goodman 2003) (Table 7.1). It is highly probable that additional species existed in the tropical east of Madagascar, adapted to a cassowary-like lifestyle, but the fossil record from these forests is entirely absent (Monnier and Lamberton 1955). Moreover, the degree of sexual dimorphism is completely unknown. Using ancient DNA sequence data, Bunce and colleagues (2003) demonstrated that underestimation of sexual dimorphism in moa has lead to an overestimation of species diversity in *Dinornis*.

Table 7.1 Species designations and associated synonyms within Aepyornithidae. *After* Hawkins and Goodman 2003.

Order Aepyornithiformes

Family Aepyornithidae

Aepyornis gracilis *Monnier, 1913*

A. hildebrandti Burckhardt, 1893

syn. A. mulleri Milne-Edwards and A. Grandidier, 1894

A. maximus I. Geoffroy-Saint-Hilaire, 1851

syn. A. modestus Milne-Edwards and A. Grandidier, 1869

syn. A. titan Andrews, 1894

syn. A. ingens Milne-Edwards and A. Grandidier, 1894

A. *medius* Milne-Edwards and A. Grandidier, 1866

syn. A. grandidieri Rowley, 1867

syn. A. cursor Milne-Edwards and A. Grandidier, 1894

syn. A. lentus Milne-Edwards and A. Grandidier, 1894

Mullerornis agilis Milne-Edwards and A. Grandidier, 1894

M. betsilei Milne-Edwards and A. Grandidier, 1894

M. grandis Lamberton, 1934

M. rudis Milne-Edwards and A. Grandidier, 1894

syn. Flacourtia rudis Andrews, 1894

Whilst *Aepyornis elephas* was the most massive bird of the Late Quaternary, its egg was the largest ever laid, even larger than any dinosaur egg (Fuller 2000). It measured about 32 cm × 24 cm and had a volume seven times that of an ostrich egg, and up to 170 times greater than a chicken egg (Cauderay 1931; Amadon 1947). It is claimed that this prodigious egg, carried back to Arabia by merchants, gave rise to the legend of the *roc* or *rukh.*

The position of *Mullerornis* has been tentatively placed among the ratites, having diverged from the cassowary, emu, kiwi and ostrich group after 78 Myr ago (Cooper et al. 2001). However, this result was derived from limited sequence data, largely a result of the paucity of well-preserved elephant bird skeletal material, a consequence of both the hot and humid environment on Madagascar and the age of the samples. By contrast to skeletal remains, elephant bird eggshell is very abundant, with some beaches seeming to be "nearly paved with broken eggshells" (Dewar 1984). In this paper we show how additional sequence data from *Aepyornis* demonstrates its affinity with *Mullerornis* as well as its position within the paleognate tree.

7.3 Sites and Sample Information

Five fragments of *Mullerornis* eggshell and thirteen fragments of *Aepyornis* eggshell were subjected to DNA extraction and amplification (Table 7.3). These ranged in age from 1415 to >30 000 years BP, and were collected from the southern tip of Madagascar (Figure 7.1). Identification of eggshell to genus level was made on the basis of their thickness. Skeletal remains indicate that *Aepyornis* was larger than *Mullerornis* (Amadon 1947), and eggshells with a thickness greater than 3-4 mm were

therefore assigned to the genus *Aepyornis*, with thinner eggshells assigned to the genus *Mullerornis* (Burney 1997).

In addition, DNA from 27 sediment samples from southern Madagascar (Figure 7.1) was extracted (Table 7.4).



Figure 7.1 Location of sample sites in Madagascar and the major rivers of the south.

7.4 Methods

7.4.1 DNA Extraction and Amplification

DNA was extracted from eggshell, purified, amplified, cloned and sequenced following the protocol described in Bunce and colleagues (in prep.). The primers used and conditions are listed in Table 7.5. Annealing temperature was 55 ºC and 60 PCR cycles used. Independent replication was carried out in the Murdoch University, Western Australia.

DNA was extracted from approximately 8g of sediment using a combined Sergey Bulat protocol (Bulat et al. 2000) and Cambio PowerMax Soil kit protocol. Chloroplast *trn*L DNA was selectively amplified using *trn*L*g* and *trn*L*h* primers (Taberlet et al. 2007) and positive results subsequently cloned and sequenced.

7.4.2 Phylogenetic Analysis

Elephant bird phylogeny was inferred using 290 12S base pairs and 273 CytB base pairs of concatenated sequences from *Aepyornis* and previously published *Mullerornis* sequences (Cooper *et al.* 2001). Bayesian inference within MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001) and maximum likelihood (ML) within PAUP*4.0b10 (Swofford, 2002) were used. In each case, substitution models were partitioned among protein-coding and RNA-coding sequences and within these, among sites for which all taxa were aligned (1. Protein, 2. RNA), sites missing *Mullerornis* only (3. Protein) and sites missing *Aepyornis* only (4. Protein, 5. RNA). Branch-lengths were separately optimised for each partition under ML, while being kept proportional with rate multipliers for the Bayesian analyses. Substitution model categories for each partition were assigned according to ModelTest 3.06 (Posada and

Crandall, 1998) AIC recommendations. In each case these were GTR+I+Γ_4 for the protein partitions and HKY+I+Γ_4 for the RNA partitions.

Monophyly of neither ratites nor palaeognathae was enforced in the Bayesian analysis. However, monophyly was enforced for several outgroup taxa (Anseriformes, Galloanserae, Charadriiformes and seabirds) as well as the tinamou-moa grouping. Each of these groupings is well resolved with larger datasets (Hackett et al. 2008; Phillips et al. in prep) and enforcing their monophyly ensures that the placement of elephant birds is consistent with known phylogenetic context. Among birds it has been shown that compositional heterogeneity influences phylogenetic reconstructions among data partitions for which saturation has greatly eroded phylogenetic signal (*e.g.* 3rd codon positions) and that RY-coding (R=A,G; Y=C,T) can ameliorate these effects (Harrison et al. 2004). However, RY-coding the terminal triplet position removes too much information from the limited elephant bird dataset, and for this reason Y-coding, in which only C and T are lumped is used as a compromise. Notably the main composition bias among ratites tends to be C vs T at protein 3rd codon positions rather than between A and G (purines R).

7.5 Results

7.5.1 Eggshell

Of the five *Mullerornis* eggshells, and the 13 *Aepyornis* eggshells analysed, three (or 60%) and seven (or 54%) respectively, produced DNA sequences identifiable as of ratite origin giving a combined success rate of 56%. However, if samples older than 3000 years are excluded, this success rate rises to 66% for both genera.

Eggshell fragments <3.00mm in width yielded only DNA sequences identical to that of *Mullerornis* (Cooper et al. 2001), and eggshell fragments >3.00mm yielded ratite DNA, presumably from *Aepyornis.*

Five elephant bird eggshells extracts (three *Aepyornis* and one *Mullerornis*) gave control region (CR) sequences, and two *Aepyornis* samples were intensively amplified producing a total of 290 12S base pairs and 273 CytB base pairs. Sequence data are presented in Table 7.2, and the results of the phylogenetic analysis are shown in Figure 7.2.

The CR is a noncoding region of approximately 1200 nucleotides spanning both sides of the arbitrary "0" position of the mt genome. The DNA sequence of the control region is termed hypervariable because it accumulates point mutations at approximately ten times the rate of nuclear DNA, and is therefore the mtDNA gene of preference for phylogeographical and population genetic studies of a variety of organisms.

7.5.2 Sediments

Of the 27 sediment samples extracted it was only possible to amplify plant DNA from six of them. Again there appears to be an age bias, with the youngest sample (EH717) producing the greatest number of sequences. Universal bird and mammal primers as well as lemur and fossa (*Cryptoprocta ferox*) primers were also applied to the extracts, but these failed to produce any meaningful results.

Table 7.5 List of 12S, CytB and *trn*L primers.

Primer Name	**Sequence**
Ratite 12S F2	5'-AACCCACGAAAGCTAAGGCA
Ratite 12S R1	5'-GTAGTTCTCRGGCGGATACT
Bird12S K2	5'-GCCTACATACCGCCGTCSC
[a]Bird12S H	5'-CCTTGACCTGTCTTGTTAGC
Ratite 12S F5	5'-CTAACAAGACAGGTCAAGGTAT
Ratite 12S R4	5'- CCTATTTTACTGCTAAATCCG
Ratite 12S G2	5'-CACCCGACCATCTCTTGCC
Bird 12S F1	5'- CGAGAACTACGAGCACAAAC
Ratite 12S ER	5'-GAACAGGCTCCTCTAGGTGGG
Ratite CytB F2	5'-CATCTCAGCCTGATGAAACTT
Ratite CytB R2	5'-GAAGTGTCGGCTGTATAGTGTAT
Ratite CytB F4	5'-ACACTTCCTTAGCCTTYTCA
Ratite CytB R4	5'-AAGAATGATGCTCCRTTTGCA
Ratite CytB F3b	5'-ATCACCCAAATTCTAACAGGAT
Ratite CytB R3b	5'-TCATCCATACTGTATGTTCCG
Ratite CytB F5	5'-GTATGGATGACTAATTCGCAAC
Ratite CytB R5b	5'-TTATAGAGGTADGAGCCRTAG
Ratite CytB F6b	5'-CGGAGATCCAGAAAACTTCAC
Ratite CytB R6b	5'-GGATAGCGTATGCGAATAGAA
Ratite CytB F7	5'-ACACATCAAACCAGAATGATAC
Ratite CytB R7	5'-GGATTAGTACGGAGGCRGC
Ratite CytB F8	5'-TATCCCTAACAAACTAGGAGG
Ratite CytB R8b	5'-CGGAAKGTTATTGARCGTTGT
Ratite CytB F11	5'-CTACACATCGGACGAGGATT
Ratite CytB R11	5'-CCTACRAAKGCAGTTGCTAT
[b]*trn*L*g*	5'-GGGCAATCCTGAGCCAA
[b]*trn*L*h*	5'-CCATTGAGTCTCTGCACCTATC

[a](Cooper et al. 1992), [b](Taberlet et al. 2007). All other primers designed during this study.

Table 7.2 Showing raw sequence data of *Aepyornis* and *Mullerornis.* ~ indicates missing bases.

Genus/Gene	**Sample**	**Sequence**
Aepyornis/12S	MAD93-A8	CAAACTGGGATTAGATACTCCACTATGCTTAGCCCTAAATCCCGATACTTCACATACACA~GCTTAAAACTCTAAGGACT TGGCGGTGCCCTAAA~TATAATCGATAACCCACGATAAACCTGACCGCTCCTCGCCAAAATGCAGCCTACATACCGCCG TCGCCAGCTCGCCTAAATGATAGAACAATAGCGAGCACAACAGCTACACC~AGCATATGAGATGGAAGAAATGGGCTA CATTTTCTAAGATAGAACATTCAACGAAAGAGARCATGAAACATCTCTCAGAAGG
Aepyornis/CytB	MAD93-A8/MAD95-12	TGGATCCTTACTAGGAATCTGTCTAATCACCCAAATTCTAACAGGATTACTACTAGCCATACACTATACAGCTGACACCT CACTAGCCTTCTCATCCGTAGCTCATACCTGCCGGAACATACAGTATGGATGACTAATTCGCAACCTCCATGCAAATGGA GCATCATTCTTCTTCATCTGTATCTACCTACACATCGGACGAGGATTTTAC~CCCAGCCAACCCCCTAGGTACACCTCCAC ATATCAAACCAGAATGATACTTTCTATTCGCATACGCTATCCTACGCTCTATCCCTAACAAACTAGGAGGAGTACTTGCC CTAGCTGCCTCCGTACTAATCCTATTCCTAATACCACTCCTACACAAATCCAA
Aepyornis/CR	MAD95-12	ACATCAATGGCTCTCGGTCTTCGTCTTTCCTGGCAGGCCGATTAATGAGACGGTTTGCGTATATGTGGAAT
Aepyornis/CR	MAD93-K13	GCATCCATGGACTTCGGGCTGCGTCTCTCCCGAGGGCCGATTAATGAGACGGTTGGCGTATATGGGGAAT
Aepyornis/CR	MAD95-47	ACATCAATGGCTCTCGGTCTTCGTCTTTCCTGGCAGGCCGATTAATGAGACGGTTTGCGTATGTGTGGAAT
Aepyornis/CR	MAD93-A8	ACATCAATGGCTCTCGGTCTTCGTCTTTCCCGGCAGGCCG
Mullerornis/CR	MAD97-26	ACATTAATGGCTCTCGGTCTTCGTCTTTCCCGGCAGGCCGATTAATGAGACGGTTTGCGTATATGGGGAAT

Table 7.3. Showing details of eggshell analyses and primers used. [a]Ratite 12S F2, Ratite 12S R1; [b]Bird12S K2, Bird 12S H; [c]Ratite 12S F5, Ratite 12S R4; [d]Ratite 12S G2, Ratite 12S R1; [e]Bird 12S F1, Ratite 12S ER; [f]Ratite CytB F2, Ratite CytB R2; [g]Ratite CytB F4, Ratite CytB R4; [h]Ratite CytB F3b, Ratite CytB R3b; [i]Ratite CytB F5, Ratite CytB R5b; [j]Ratite CytB F6b, Ratite CytB R6b; [k]Ratite CytB F7, Ratite CytB R7; [l]Ratite CytB F8, Ratite CytB R8b; [m]Ratite CytB F11, R11; [x]Ratite CR F1, Ratite CR R1; [y]Ratite CR F1, Ratite CR F2, (Superscript capital letters indicate independent replication).

Sample I.D.	**Eggshell Type**	**Site Name**	**Site Coordinates**	**Uncal RC Age (years BP)**	**DNA Results**
MAD93-A2	*Aepyornis*	Maroaloke	25°13'S, 46°12'E	2285±40	*Did not work*
MAD93-A7	*Aepyornis*	Maroaloke	25°13'S, 46°12'E	>30k	*Did not work*
MAD93-A8	*Aepyornis*	Maroaloke	25°13'S, 46°12'E	1940±45	*Aepyornis*[a,f,g,h,i,j,k,l,y]
MAD93-A10	*Aepyornis*	Maroaloke	25°13'S, 46°12'E	1415±40	*Aepyornis*?
MAD93-K12b	*Aepyornis*	North of Talaky	25°23'S, 45°50'E	Pleistocene	*Did not work*
MAD93-K13b	*Aepyornis*	Maroaloke	25°13'S, 46°12'E		Ratite[x,y]
MAD95-12	*Aepyornis*	Cap St Marie	25°36'S, 45°14'E		*Aepyornis*[m,x,y]
MAD95-47	*Aepyornis*	Meandrare estuary	25°09'S, 46°26'E		*Aepyornis*[C]
MAD95-49	*Aepyornis*	Meandrare estuary	25°09'S, 46°26'E		*Aepyornis*[a,b,c,C,d,e]
MAD95-50A	*Mullerornis*	Meandrare estuary	25°09'S, 46°26'E		*Mullerornis*[a]
MAD97-15	*Mullerornis*	Talaky	25°27'S, 45°44'E		*Did not work*
MAD97-19	*Aepyornis*	Talaky	25°27'S, 45°44'E	4030+/-40	*Did not work*
MAD97-23	*Mullerornis*	Talaky	25°27'S, 45°44'E		*Did not work*
MAD97-26	*Mullerornis*	Talaky	25°27'S, 45°44'E		*Mullerornis*[a,x]
MAD97-45	*Aepyornis*	Talaky	25°27'S, 45°44'E	2285±35	*Did not work*
MAD97-76	*Mullerornis*	Talaky	25°27'S, 45°44'E		*Mullerornis*[y]
MAD00-50A	*Aepyornis*	Reampeko	25°16'S, 44°39'E	1442±29	*Aepyornis*[x]

Table 7.4. Showing results of plant DNA from sediments

	Sample I.D.	Sample Type	Site Name	**Site Coordinates**	**Approximate age**	**DNA Results**
EH763	MAD95-6	Dune sand	Cap St Marie	25°36'S, 45°14'E	Holocene	*Did not work*
EH850	MAD95-6	Dune sand	Cap St Marie	25°36'S, 45°14'E	Holocene	*Did not work*
EH764	MAD95-44	Dune sand	Meandrare estuary	25°09'S, 46°27'E	Holocene	*Did not work*
EH765	MAD97-14	Dune sand	Talaky	25°27'S, 45°43'E	Holocene	*Did not work*
EH766	MAD97-17	Dune sand	Talaky	25°27'S, 45°43'E	Holocene	*Did not work*
EH851	MAD97-17	Dune sand	Talaky	25°27'S, 45°43'E	Holocene	*Did not work*
EH767	MAD97-18	Dune sand	Talaky	25°27'S, 45°43'E	Holocene	*Did not work*
EH768	MAD97-20	Dune sand	Talaky	25°27'S, 45°43'E	Holocene	*Did not work*
EH769	MAD97-21	Dune sand	Talaky	25°27'S, 45°43'E	Holocene	*Did not work*
EH853	MAD97-29	Dune sand	Talaky	25°27'S, 45°43'E	Holocene	*Did not work*
EH713	MAD97-34	Cemented sand	Talaky	25°27'S, 45°43'E	Pleistocene	Bryophyta, Lauraceae
EH714	MAD97-50	Dune sand	Talaky	25°28'S, 45°42'E	Holocene	*Did not work*
EH770	MAD97-52	Dune sand	Talaky	25°28'S, 45°42'E	Holocene	*Did not work*
EH852	MAD97-52	Dune sand	Talaky	25°28'S, 45°42'E	Holocene	*Did not work*
EH854	MAD97-59	Organic estuarine sand	Talaky	25°28'S, 45°42'E	Holocene	*Did not work*
EH715	MAD97-62	Estuarine Sand	Talaky	25°28'S, 45°42'E	Holocene	Solanaceae, Poaceae
EH716	MAD97-66	Dune sand	Talaky	25°28'S, 45°40'E	Holocene ?	*Did not work*
EH717	MAD97-79	Dune sand below buried soil.	Talaky	25°28'S, 45°41'E	Late Holocene-Medieval	Cupressaceae, Sapotaceae, Rutiaceae, Euphorbiaceae, *Tetracena madagascarensis, Thylachium africanum, Plumbago zeylanica*.
EH771	MAD97-79	Dune sand	Talaky	25°28'S, 45°41'E	Holocene	*Did not work*
EH709	MAD00-24	Sand above organic sediment	Anovao	25°17'S, 44°37'E	Holocene ?	Bryophyta, Crotoneae, Spinaceae/Ochromondaceae, Meliaceae
EH772	MAD00-25	Sand below organic sediment	Anovao	25°17'S, 44°37'E	Holocene ?	*Did not work*
EH710	MAD00-25	Sand below organic sediment	Anovao	25°17'S, 44°37'E	Holocene ?	*Did not work*
EH773	MAD00-38	Dune sand	Anovao	25°18'S, 44°39'E	Holocene	*Did not work*
EH774	MAD00-76	Dune sand		25°04'S, 44°06'E		*Did not work*
EH711	MAD00-86	Sandy clay	Itampolo	24°40'S, 43°57'E	Holocene ?	Bryophyta
EH712	MAD00-89	Peaty Sand	Itampolo	24°40'S, 43°57'E	Holocene ?	Bryophyta
EH856	MAD00-89	Peaty Sand	Itampolo	24°40'S, 43°57'E	Holocene ?	*Did not work*

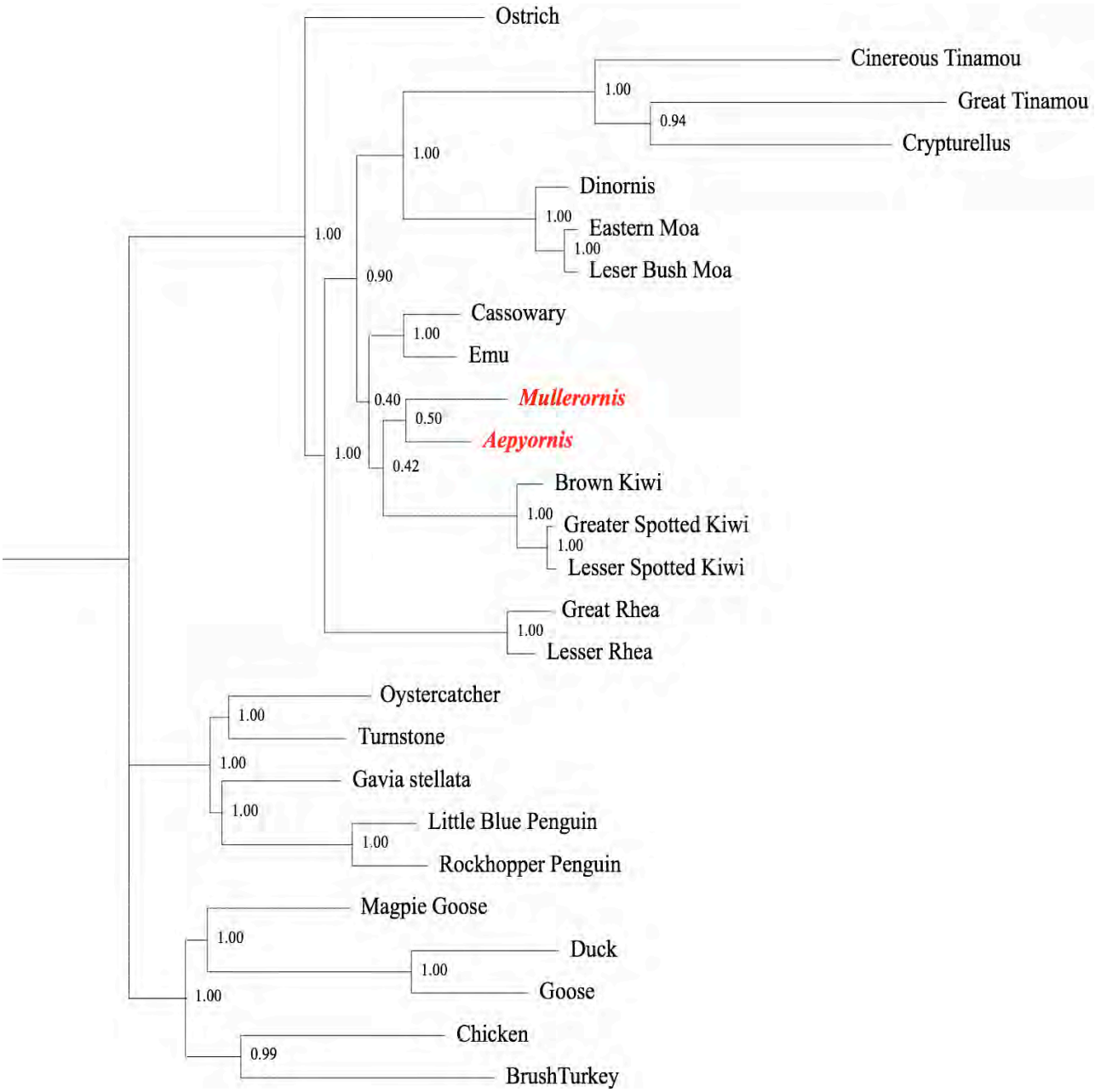


Figure 7.2 Bayesian Inference Tree showing the position of *Mullerornis* and *Aepyornis*. Bayesian Posterior Values are given for nodes.

7.6 Discussion

7.6.1 Inter-island Phylogeography and Phylogenetics

Krause (Krause et al. 1997) wrote that:

The biogeographic origins of the modern terrestrial and freshwater vertebrate fauna of Madagascar remain one of the greatest unsolved mysteries of natural history.

Due to the high level of endemism of its flora and fauna, Madagascar has undoubtedly been isolated for a long time, but affinities to African, Asian, and Gondwanan biotas has led to speculation as to when, how, and from where various groups got to the island. In the late nineteenth and early twentieth centuries land bridges between India, South America, Madagascar and Africa were invoked, but today plate tectonic theory holds sway.

Current knowledge suggests that direct continental connection between Madagascar and Africa was severed at about 160-150 Myr ago (Krause et al. 1997), although a circuitous route from Africa to Madagascar via South America, Antarctica and the Indian subcontinent might have been open until about 130-120 Ma (Smith et al. 1994). During the Cretaceous, Madagascar remained physically and biotically connected to the Indian sub-continent, although by this stage, which lasted from about 120 to about 85 Myr ago, both were fully isolated from other Gondwanan landmasses. From the late Cretaceous, to the present, Madagascar has been separated from the Indian subcontinent, its vertebrate fauna evolving in isolation, with occasional over-water immigration primarily from Africa across the Mozambique Channel.

The results of the phylogenetic analysis suggest that elephant birds are part of the Australian/NZ/Antarctic radiation and got to Madagascar from there - probably with the aid of flight and maybe an emergent Kerguelen plateau. In spite of the current proximity of Madagascar to Africa, elephant birds are not closely relate to ostriches, and appear to derive from among the Casuariiformes-kiwi-tinamous-moa clade. Elephant birds in fact group most closely with Kiwi, which given the difference in size might at first seem surprising. However, size plasticity among island birds is well established (Bunce et al. 2005), and it is worth noting that the kiwi lays the largest egg per body size of any living bird (Calder 1979).

The monophyly of the elephant birds is evident from the data, although the support is low, suggesting that these two genera are more distantly related than previously thought.

Although the results of the phylogenetic analysis would be greatly strengthened by the addition of further sequence data, this paper demonstrates that aDNA can survive in eggshells, even when *Aepyornis* bones of equivalent age consistently fail to produce positive results (Cooper, A. *pers. comm.*).

7.6.2 Intra-island Phylogeography and Phylogenetics

Whilst the phylogenetic relationships and evolutionary history of the aepyornithids across the far-flung remnants of Gondwana is fascinating, of equal importance is the possibility that aDNA from eggshell allows the investigation intra-island phylogenetics and phylogeography of elephant birds. The systematics of elephant birds is in disarray, and whilst geographic barriers have been shown to create genetic isolation and distinct zones of species distribution, for example in lemurs (Pastorini et al. 2003), the nature of aepyornithid biogeography within Madagascar is completely

unknown. Furthermore, as ratite eggshell is particularly amenable to radiocarbon dating (Freundlich *et al.* 1989; Higham 1994; Vogel *et al.* 2001; Holdaway *et al.* 2002), it may well be possible in the future to track the movement of different species of elephant bird (or haplotypes) across time as well as space.

Preliminary results of *Aepyornis* control region mtDNA offers a tantalising glimpse at inter-individual variation, with three of the four sequences generated being distinct from one another (with the fourth being a very short sequence). The eggshells from which these sequences were produced came from different sites which are geographically isolated by rivers (Figure 1). Obviously without the generation of substantially more sequences from eggshells from other parts of the island, insights into phylogeographic patterns and inter-species variation is not possible, but this result clearly demonstrates the very real potential of this approach.

A better phylogeographic understanding would, in turn, allow the investigation of the nature of human impact on Madagascan elephant bird populations. There is evidence of predation of elephant birds by the first Malagasy, with remains of eggshell having been found in archaeological deposits (Dewar 1984; Dewar and Rakotovololona 1992). In contrast to the extinction of the moa, however, which may have occurred in less than 100 years after Polynesian settlement (Holdaway and Jacomb 2000), there is a gap of over 1000 years between the earliest date for a human presence (2325 ± 43 yr BP) (Burney et al. 2004) and the most recent date from an *Aepyornis* eggshell 1000 ± 150 yr BP (Berger et al. 1975). Also, whereas overkill (perhaps combined with habitat loss) is the generally accepted mechanism for the extinction of the moa in New Zealand, (Anderson 1989; Holdaway and Jacomb 2000) several other additional causes have been suggested for the extinction of the elephant birds in Madagascar:

hyperdisease (MacPhee and Marx 1997), introduction of competitors (Dewar 1984), climate change or desiccation (Mahé and Sourdat 1972), and anthropogenic fire (Morat 1973). Whatever the cause, or combination of causes of elephant bird extinction, it is aDNA from eggshell that will help to provide an improved understanding of this process.

7.6.3 Ancient DNA in Eggshell

Whilst there has been some speculation about location of aDNA in bone (Salamon et al. 2005), eggshell, despite also being calcareous (calcium carbonate in the form of the mineral calcite), is a very different biomaterial; the distribution of aDNA may also be very different. Furthermore, in the case of ratite eggshells, there is the possibility that the DNA is derived from the mother, the chick, or possibly even a male incubating bird. This last possibility is perhaps the least likely because shell cuticle is impermeable, although it is punctuated by fine pores that permit entry and exit of atmospheric gases, including water vapour (Burley and Vadehra 1989).

Although more than 97% of eggshell is inorganic material, the remainder consists of an organic matrix that contains about 11% of the shell protein. This matrix is not distributed evenly with a plug concentrated at the mammillary core. Just inside the shell are two membranes: a thick outer membrane (the shell membrane), and a thin inner membrane (the egg membrane) (Figure 7.3) (Burley and Vadehra 1989). These membranes are 0.10 and 0.03 mm thick, respectively, in the ostrich (Deeming 1997). Shell-membrane fibres extend into the mammillary core and attach the membrane to the calcified layer (Robinson and King 1968; Bellairs and Boyde 1969), and this area of the eggshell could also be the site of greatest DNA density.

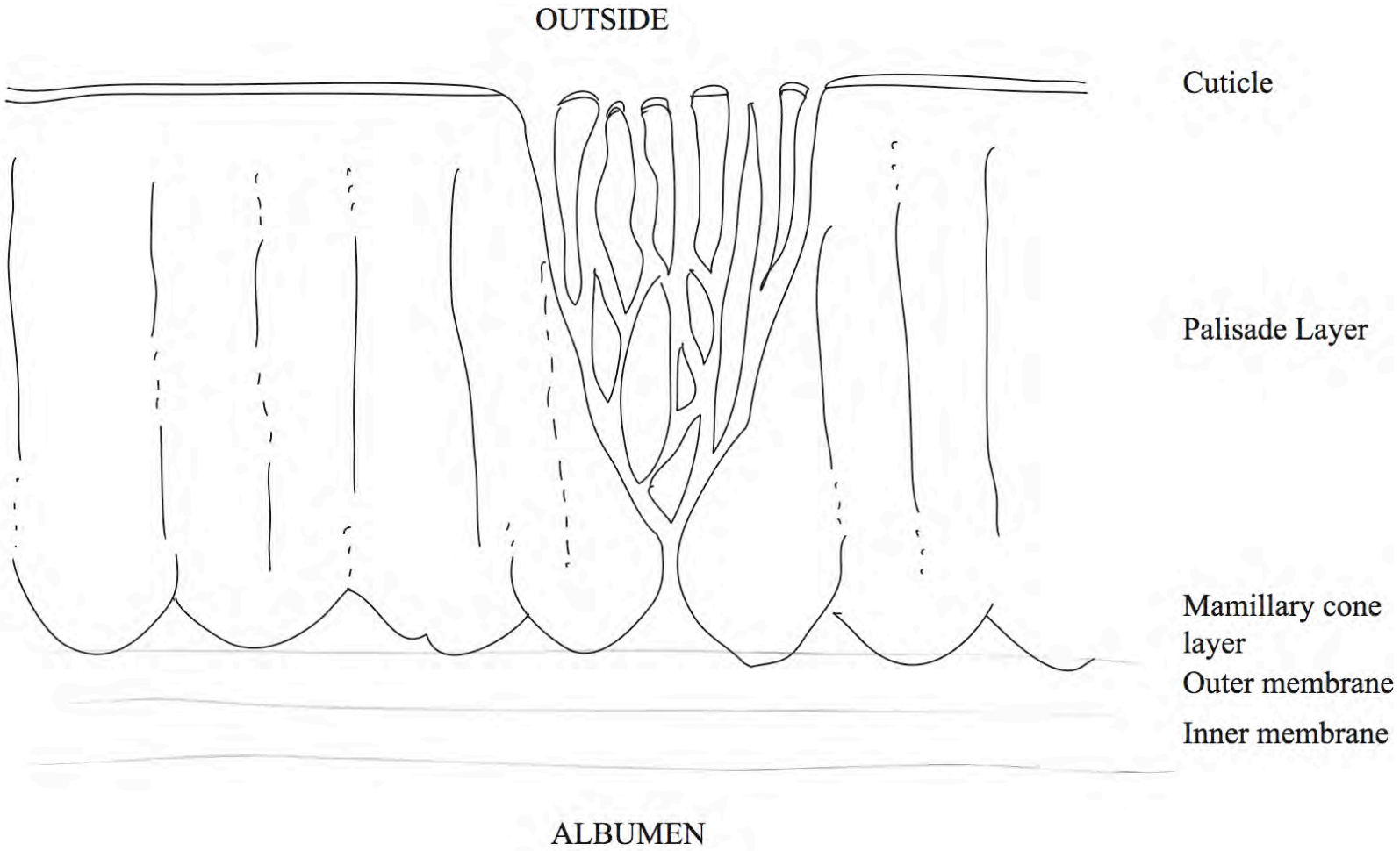


Figure 7.3. Radial cross section of a stylised ostrich eggshell (after Deeming 1997 and Romanoff 1949).

The activity of bacterial exonucleases is thought to play an important role in the post-mortem degradation of nucleic acids. However, the shell membranes contain antibacterial enzymes such as lysozmes (Vadehra et al. 1972) and β-*N*-acetylglucosaminidase (Winn and Ball 1975), which would help to limit bacterial activity on the inner shell surface, at least in the first moments after hatching and until desiccation occurs which would itself reduce the activity of bacteria and their exonucleases.

7.6.4 Ancient DNA from Sediments

By contrast with the eggshell results, the results from the sediments were disappointing. No mammal or bird DNA was amplified, and the success rate of plant DNA amplification was comparatively low. The sites where the sediment samples were sourced are all open to the elements, and the low success rate is perhaps a reflection of the rate of DNA degradation in these hot and seasonally wet environments.

Although the plant results come from families, genera and species found today in Madagascar, the degree of resolution to species level is poor for the reason that the *trn*L database on GenBank of tropical plants lags behind that of arctic plants (*cf* Chapter 4). Furthermore, until vertical movement of DNA in open, non-frozen locations is better understood, any palaeological interpretation of the plant results is open to question.

7.7 Conclusions

As a novel source of ancient DNA, eggshell has an enormous potential for revealing otherwise hidden aspects of elephant bird phylogeny, and phylogeography, as well as addressing the impacts of climate change and human exploitation in Madagascar. This potential is even more marked in Madagascar where skeletal remains of elephant birds are rare, and in the tropical east, non-existent, and what remains that do exist are poor sources of aDNA. By contrast, elephant bird eggshells are common, and are thus the material of choice for future studies of ratites in this environment.

Chapter 8

Conclusions and Future Directions

Now this is not the end.
It is not even the beginning of the end.
But it is, perhaps, the end of the beginning.

Winston Churchill 10 November 1942

In each of my five core chapters I hope to have demonstrated that I have advanced the field of aDNA through the development of new applications for ancient sedimentary DNA, a novel use for aDNA from skeletal material, and also the potential of aDNA research from "non-traditional" materials, by which I mean substrates other than bone or teeth, which have up to now been the mainstay of aDNA research.

One of the drawbacks of *sed*aDNA is that sufficient phylogenetic information for the species in question may not be contained within a short section of DNA (typically less than 150 bp). It is not advisable to concatenate overlapping sequences, as one can with aDNA extracted from tissues, as it is impossible to know how many individuals have contributed to the DNA recovered from the sediments. For the same reason, the degree of damage of the samples cannot be determined as even if the amplicons are cloned, sequence variation might be due to inter-individual variation, and not *post mortem* damage. Notwithstanding this handicap, the use of *sed*aDNA as an indicator for the presence of certain species has the potential to revolutionise the fields of paleobiology, paleoarchaeology, climatology, and extinction biology.

The use of sedaDNA to provide a more refined estimate for the last appearance dates of arctic megafauna will enable a better understanding of the process of their extinction; of the relative contributions of human pressure and climate change. Past climate change itself can now be revealed, at the local level, now that the groundwork has been laid whereby the relationship between modern plant distributions and abundance, and modern sedimentary plant DNA has been established. Further work to correlate ancient plant DNA with pollen data, plant macrofossil remains, and beetle fauna is currently underway. Such knowledge will enable models to be built which

can more accurately predict the impact of future climate change on the most thermally vulnerable of ecosystems, the arctic.

However, in order for the relatively new discipline of *sed*aDNA to become firmly entrenched within the pantheon of paleo-techniques, a better understanding of the process by which aDNA persists in sediments is required. A series of experiments in Svalbard involving bacterial DNA, faecally and urine-derived horse DNA, qPCR, molecular staining and other techniques is being devised for the summer of 2009 and a collaborator from the University of Bournemouth is sampling sediments from around a deer carcass in a temperate environment. As well as examining the persistence of DNA in sediments in these sites, the vertical and horizontal movement of this molecule will also be studied.

Whilst there are several studies that suggest that in permafrost environments, and also in temperate environments *in the absence of free water* the migration of aDNA in sediments is curtailed, when these conditions are not met, the temporal resolution of this technique is questionable. The longevity of free DNA is thought to be brief, but if non-contemporary free DNA can displace contemporary sediment-bound aDNA, then only in dry environments will this discipline have a future. Only further studies can reveal how widespread is this phenomenon.

More universal is the application of aDNA provenancing to determine the origin of plants and animals which show strong phylogeographic signals, and for which a genetic database has already been compiled. Provenancing has a long-established history within the field of archaeology, especially for pottery, ceramics and stone

tools, and now aDNA provenancing on animal remains opens up an additional source of information, that may be of particular relevance for the study of New World, or Oceanic Island sites, where the geographical source of introduced animals is unclear.

Finally the novel use of ancient eggshells as an aDNA resource is extremely exciting. The information gained from this material can be used in the same way as that from bones: to construct phylogenies, for phylogeographical reconstructions, population structure analyses, dietary reconstruction from amino acid studies, and for extinction patterns. In Madagascar elephant bird skeletal remains are extremely rare, whereas eggshells are relatively common, and in addition, eggshell appears to preserve aDNA better than bone: several hundred bp long nuclear DNA sequences have been amplified from Moa eggshell, and a 25 000 year old Emu eggshell from Australia has also yielded DNA! The work on Madagascan elephant bird eggshell is on-going, and with the acquisition of samples from different regions on the island, a better picture of the birds' phylogeography will develop.

Another overlooked source of aDNA is that from other calcareous materials such as ancient mollusc shells and crustacean shells. Preliminary work has demonstrated that these materials do indeed yield DNA, and applications of this finding are being developed.

Literature Cited

Agnelli, A., J. Ascher, G. Corti, M. Ceccherini, P. Nannipieri and G. Pietramellara (2004). "Distribution of microbial communities in a forest soil profile investigated by microbial biomass, soil respiration and DGGE of total and extracellular DNA." Soil Biology and Biochemistry **36** (5): 859-868.

Aitken, M. (1998). An Introduction to Optical Dating, Oxford University Press, Oxford.

Amadon, D. (1947). "An estimated weight of the largest known bird." The Condor **49**: 159-164.

An C-C, Li Y and Zhu Y-X (1995). "Molecular cloning and sequencing of the 18S rDNA from specialized dinosaur egg fossil found in Xixia Henan, China." Acta Sci Nat Univ Pekinensis **31**: 140-147.

Anderson, A. (1989). "Mechanics of overkill in the extinction of New Zealand moas." Journal of Archaeological Science **16** (2): 137-151.

Anderson, A. (2003). "Investigating Early Settlement on Lord Howe Island." Australian Archaeology **57**: 98-102.

Arnold, L. J., R. G. Roberts, R. D. E. MacPhee, E. Willerslev, A. N. Tikhonov and F. Brock (2008). "Optical dating of perennially frozen deposits associated with preserved ancient plant and animal DNA in north-central Siberia." Quaternary Geochronology **3** (1-2): 114-136.

Arumuganathan, K. and E. Earle (1991). "Nuclear DNA content of some important plant species." Plant Molecular Biology Reporter **9**: 211-215.

Asara, M. (2007). "Protein sequences from mastodon and Tyrannosaurus rex revealed by mass spectrometry." Science **316** (5832): 1698.

Asher, J., M. Ceccherini, P. Nannipieri and G. Pietramellara (2005). Extracellular DNA Rise up in Soil by Water Capillarity. Proceeding of the 8th Symposium on Bacterial genetcis and Ecology (BAGECO), Lyon, France.

Austin, J. J., E. Nicholas Arnold and R. Bour (2002). "The Provenance of Type Specimens of Extinct Mascarene Island Giant Tortoises (Cylindraspis) Revealed by Ancient Mitochondrial DNA Sequences." Journal of Herpetology **36** (2): 280-285.

Banerjee, M. and T. A. Brown (2004). "Non-random DNA damage resulting from heat treatment: implications for sequence analysis of ancient DNA." Journal of Archaeological Science **31** (1): 59-63.

Barnes, I., B. Shapiro, A. Lister, T. Kuznetsova, A. Sher, D. Guthrie and M. G. Thomas (2007). "Genetic structure and extinction of the woolly mammoth, Mammuthus primigenius." Current Biology **17** (12): 1072-1075.

Barnett, R., I. Barnes, M. J. Phillips, L. D. Martin, C. R. Harington, J. A. Leonard and A. Cooper (2005). "Evolution of the extinct Sabretooths and the American cheetah-like cat." Current Biology **15** (15): R589-R590.

Barnett, R., N. Yamaguchi, B. Shapiro and V. Nijman (2007). "Using ancient DNA techniques to identify the origin of unprovenanced museum specimens, as illustrated by the identification of a 19th century lion from Amsterdam." Contributions to Zoology **76** (2): 87-94.

Bellairs, R. and A. Boyde (1969). "Scanning electron microscopy of the shell membranes of the hen's egg." Zeitschrift für Zellforschung und mikroskopische Anatomie **96**: 237-249.
Bennett, K. D. and L. Parducci (2006). "DNA from pollen: principles and potential." Holocene **16** (8): 1031-1034.
Benson, D., I. Karsch-Mizrachi, D. Lipman, J. Ostell, B. Rapp and D. Wheeler (2000). "GenBank." Nucleic Acids Research **1;28** (1): 15-18.
Berger, R., K. Ducote, K. Robinson and H. Walter (1975). "Radiocarbon Date for Largest Extinct Bird." Nature **258** (5537): 709-709.
Bertolla, F., F. van Gijsegem, X. Nesme and P. Simonet (1997). "Conditions for natural transformation of *Ralstonia solanacearum*." Applied and Environmental Microbiology **63**: 4965-4968.
Bigelow, N., L. Brubaker, M. Edwards, S. Harrison, I. Prentice, P. Anderson, A. Andreev, P. Bartlein, T. Christensen, W. Cramer, J. Kaplan, A. Lozhkin, N. Matveyeva, D. Murray, A. McGuire, V. Razzhivin, J. Ritchie, B. Smith, D. Walker, K. Gajewski, V. Wolf, Y. Igarashi, K. Kremenetskii, A. Paus, M. Pisaric and V. Volkova (2003). "Climate change and Arctic ecosystems: 1. Vegetation changes north of 55°N between the last glacial maximum, mid-Holocene, and present." Journal of Geophysical Research **108** (D19): 8170.
Binladen, J., M. T. Gilbert, J. P. Bollback, F. Panitz, C. Bendixen, R. Nielsen and E. Willerslev (2007). "The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing." PLoS ONE **2**: e197.
Blair, D., B. McDonald, A. McMahon, D. Tikel, M. Waycott and H. Marsh (in prep.). "Phylogeography and population genetic structure of the dugong in Australian waters: implications for conservation."
Blum, S., M. Lorenz and W. Wackernagel (1997). "Mechanism of retarded DNA degradation and prokaryotic origin of DNases in nonsterile soils." Systematic and Applied Microbiology **20**: 513-521.
Blum, S. A. E., M. G. Lorenz and W. Wackernagel (1997). "Mechanism of retarded DNA degradation and prokaryotic origin of DNases in nonsterile soils." Systematic and Applied Microbiology **20** (4): 513-521.
Bråthen, K.-A. and O. Hagberg (2004). "More efficient estimation of plant biomass." Journal of Vegetation Science **15**: 653-660.
Brodkorb, P. (1963). "Catalogue of fossil birds. Part 1 (Archaeopterygiformes through Ardeiformes)." Bulletin of the Florida State Museum **7**: 179-293.
Bulat, S., M. Lubeck, I. Alekhina, F. Jensen, I. Knudsen and P. Lubeck (2000). "Identification of a universally primed-PCR-derived sequence-characterized amplified region marker for an antagonistic strain of *Clonostachys rosea* and development of a strain-specific PCR detection assay." Applied and Environmental Microbiology **66** (11): 4758-63.
Bunce, M., M. Szulkin, H. Lerner, I. Barnes, B. Shapiro, A. Cooper and R. Holdaway (2005). "Ancient DNA Provides New Insights into the Evolutionary History of New Zealand's Extinct Giant Eagle." PLoS Biology **3** (1e9).
Bunce, M., T. Worthy, T. Ford, W. Hoppitt, E. Willerslev, A. Drummond and A. Cooper (2003). "Extreme reversed sexual size dimorphism in the extinct New Zealand moa Dinornis." Nature **425**: 172-175.
Burley, R. and D. Vadehra (1989). The Avian Egg, John Wiley & Sons.
Burney, D. (1997). Theories and facts regarding Holocene environmental change before and after human colonization. Natural Change and Human Impact in

Madagascar. S. M. Goodman and B. Patterson. Washington, Smithsonian Institution Press**:** 75-89.
Burney, D. A., L. P. Burney, L. R. Godfrey, W. L. Jungers, S. M. Goodman, H. T. Wright and A. J. T. Jull (2004). "A chronology for late prehistoric Madagascar." Journal of Human Evolution **47** (1-2): 25-63.
Cai, P., Q. Y. Huang and X. W. Zhang (2006). "Interactions of DNA with clay minerals and soil colloidal particles and protection against degradation by DNase." Environmental Science and Technology **40** (9): 2971-2976.
Calder, W. A. (1979). "Kiwi and Egg Design - Evolution as a Package Deal." Bioscience **29** (8): 461-467.
Cano, R., H. Poinar, N. Pieniazek, A. Acra and J. G. Poinar (1993). "Amplification and sequencing of DNA from a 120-135-million-year-old weevil." Nature **363**: 536-538.
Cano, R., H. Poinar and J. G. Poinar (1992). "Isolation and partial characterisation of DNA from the bee *Problebeia dominica* (Apidae: Hymenoptera) in 25-40 million year old amber." Medical Science Research **20**: 249-251.
Cathey, H. (1990). The 2003 US National Arboretum "Web Version" of the USDA Plant Hardiness Zone Map, USDA Miscellaneous Publication No. 1475.
Cauderay, H. (1931). "Etude sur l'Aepyornis." L'Oiseau et la Revue Française d'Ornithologie **1**: 624-44.
Cheke, A. (1987). An ecological history of the Mascarene Islands, with particular reference to extinctions and introductions of land vertebrates. Studies of Mascarene Island Birds. A. Diamond. Cambridge, Cambridge University Press**:** 5-89.
Cheke, A. and J. P. Hume (2008). Lost Land of the Dodo. An ecological history of Mauritius, Réunion & Rodrigues
. London, A & C Black.
Chiter, A., J. M. Forbes and G. E. Blair (2000). "DNA stability in plant tissues: implications for the possible transfer of genes from genetically modified food." FEBS Letters **481** (2): 164-168.
Clarke, S. J., G. H. Miller, M. L. Fogel, A. R. Chivas and C. V. Murray-Wallace (2006). "The amino acid and stable isotope biogeochemistry of elephant bird (Aepyornis) eggshells from southern Madagascar." Quaternary Science Reviews **25** (17-18): 2343-2356.
Colinvaux, P. A. (1964). "Origin of Ice Ages - Pollen Evidence from Arctic Alaska." Science **145** (363): 707-708.
Cooper, A., C. Lalueza-Fox, S. Anderson, A. Rambaut, J. Austin and R. Ward (2001). "Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution." Nature **409** (6821): 704-7.
Cooper, A., C. Mourer-Chauvire, G. K. Chambers, A. von Haeseler, A. C. Wilson and S. Paabo (1992). "Independent origins of New Zealand moas and kiwis." Proceedings of the National Academy of Sciences of the United States of America **89** (18): 8741-4.
Cooper, A. and H. Poinar (2000). "Ancient DNA: Do it right or not at all." Science **289**: 1139.
Crecchio, C. and G. Stotzky (1998). "Binding of DNA on humic acids: effect on transformation of *Bacillus subtilis* and resistance to DNase." Soil Biology and Biochemistry **30** (8-9): 1061-1067.
Crichton, M. (1991). Jurassic Park, Ballantine Books.

Czerepanov, S. (1995). Vascular plants of Russia and adjacent states (the former USSR). New York, Cambridge University Press.
de Vries, J., M. Heine, K. Harms and W. Wackernagel (2003). "Spread of Recombinant DNA by Roots and Pollen of Transgenic Potato Plants, Identified by Highly Specific Biomonitoring Using Natural Transformation of an *Acinetobacter* sp." Applied and Environmental Microbiology **69** (8): 4455-4462.
Deeming, D. (1997). Ratite Egg Incubation: A Practical Guide, Ratite Conference.
Deguilloux, M. F., M. H. Pemonge and R. J. Petit (2002). "Novel perspectives in wood certification and forensics: dry wood as a source of DNA." Proceedings of the Royal Society B: Biological Sciences **269** (1495): 1039-46.
del Pozo, O. and E. Lam (1998). "Caspases and programmed cell death in the hypersensitive response of plants to pathogens." Current Biology **8** (20): 1129-1132.
DeLong, S. B. and L. J. Arnold (2007). "Dating alluvial deposits with optically stimulated luminescence, AMS 14C and cosmogenic techniques, western Transverse Ranges, California, USA." Quaternary Geochronology **2** (1-4): 129-136.
Demanèche, S., L. Jocteur-Monrozier, H. Quiquampoix and P. Simonet (2001). "Evaluation of biological and physical protection against nuclease degradation of clay-bound plasmid DNA." Applied and Environmental Microbiology **67** (1): 293-299.
Demuro, M., R. G. Roberts, D. G. Froese, L. J. Arnold, F. Brock and C. B. Ramsey (in press). "Optically stimulated luminescence dating of single and multiple grains of quartz from perennially frozen loess in western Yukon Territory, Canada: Comparison with radiocarbon chronologies for the late Pleistocene Dawson tephra." Quaternary Geochronology **In Press, Corrected Proof**: 136.
DeSalle, R., M. Barcia and C. Wray (1993). "PCR jumping in clones of 30-million-year-old DNA fragments from amber preserved termites (Mastotermes electrodominicus)." Experientia **49** (10): 906-9.
DeSalle, R., J. Gatesy, W. Wheeler and D. Grimaldi (1992). "DNA sequences from a fossil termite in Oligo-Miocene amber and their phylogenetic implications." Science **257** (5078): 1933-6.
Dewar, R. (1984). Extinctions in Madagascar: The Loss of the Subfossil Fauna. Quaternary extinctions: A prehistoric revolution, Tuscon: University of Arizona Press.
Dewar, R. and S. Rakotovololona (1992). "La chasse aux subfossiles: les preuves du onzième siècle au treizième siècle." Taloha **11**: 4-15.
Diaz-Ravina, M., A. Frostegaard and E. Baaaaath (1994). "Thymidine, leucine and acetate incorporation into soil bacterial assembalges at different temperatures." FEMS Microbiology Ecology **14**: 221-232.
Elven, R. (2007). Checklist of the Panarctic Flora (PAF) Vascular Plants.
Fish, S. A., T. J. Shepherd, T. J. McGenity and W. D. Grant (2002). "Recovery of 16S ribosomal RNA gene fragments from ancient halite (vol 417, pg 432, 2002)." Nature **420** (6912): 202-202.
Flury, M., H. Fluhler, W. A. Jury and J. Leuenberger (1994). "Susceptibility of Soils to Preferential Flow of Water - a Field-Study." Water Resources Research **30** (7): 1945-1954.

Fontaine, S., S. Barot, P. Barre, N. Bdioui, B. Mary and C. Rumpel (2007). "Stability of organic carbon in deep soil layers controlled by fresh carbon supply." Nature **450** (7167): 277-80.

Freundlich, J. C., R. Kuper, P. Breunig and H. G. Bertram (1989). "Radiocarbon Dating of Ostrich Eggshells." Radiocarbon **31** (3): 1030-1034.

Froggatt, P. and D. Lowe (1990). "A review of late Quaternary silicic and some other formations from New Zealand: their stratigraphy, nomenclature, distribution, volume, and age." New Zealand Journal of Geology and Geophysics **33**: 89-109.

Fuller, E. (2000). Extinct Birds, Oxford University Press.

Gallori, E., M. Bazzicalupo, L. Dal Canto, R. Fani, P. Nannipieri, C. Vettori and G. Stotzky (1994). "Transformation of *Bacillus subtilis* by DNA bound on clay in non-sterile soil." FEMS Microbiology Ecology **15**: 119-126.

Gebhard, F. and K. Smalla (1999). "Monitoring field releases of genetically modified sugar beets for persistence of transgenic plant DNA and horizontal gene transfer." FEMS Microbiology Ecology **28** (3): 261-272.

Gilbert, M., A. Hansen, E. Willerslev, L. Rudbeck, I. Barnes, N. Lynnerup and A. Cooper (2003). "Characterization of genetic miscoding lesions caused by postmortem damage." American Journal of Human Genetics **72** (1): 48-61.

Gilbert, M. T., L. P. Tomsho, S. Rendulic, M. Packard, D. I. Drautz, A. Sher, A. Tikhonov, L. Dalen, T. Kuznetsova, P. Kosintsev, P. F. Campos, T. Higham, M. J. Collins, A. S. Wilson, F. Shidlovskiy, B. Buigues, P. G. Ericson, M. Germonpre, A. Gotherstrom, P. Iacumin, V. Nikolaev, M. Nowak-Kemp, E. Willerslev, J. R. Knight, G. P. Irzyk, C. S. Perbost, K. M. Fredrikson, T. T. Harkins, S. Sheridan, W. Miller and S. C. Schuster (2007). "Whole-genome shotgun sequencing of mitochondria from ancient hair shafts." Science **317** (5846): 1927-30.

Gilbert, M. T., E. Willerslev, A. J. Hansen, I. Barnes, L. Rudbeck, N. Lynnerup and A. Cooper (2003). "Distribution patterns of postmortem damage in human mitochondrial DNA." American Journal of Human Genetics **72** (1): 32-47.

Gilbert, M. T. P., D. I. Drautz, A. M. Lesk, S. Y. W. Ho, J. Qi, A. Ratan, C.-H. Hsu, A. Sher, L. Dalen, A. Gotherstrom, L. P. Tomsho, S. Rendulic, M. Packard, P. F. Campos, T. V. Kuznetsova, F. Shidlovskiy, A. Tikhonov, E. Willerslev, P. Iacumin, B. Buigues, P. G. P. Ericson, M. Germonpre, P. Kosintsev, V. Nikolaev, M. Nowak-Kemp, J. R. Knight, G. P. Irzyk, C. S. Perbost, K. M. Fredrikson, T. T. Harkins, S. Sheridan, W. Miller and S. C. Schuster (2008). "Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes." Proceedings of the National Academy of Sciences **105** (24): 8327-8332.

Gilbert, M. T. P., B. Shapiro, A. Drummond and A. Cooper (2005). "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** (7): 1053-1060.

Gilbert, M. T. P., A. S. Wilson, M. Bunce, A. J. Hansen, E. Willerslev, B. Shapiro, T. F. G. Higham, M. P. Richards, T. C. O'Connell, D. J. Tobin, R. C. Janaway and A. Cooper (2004). "Ancient mitochondrial DNA from hair." Current Biology **14** (12): R463-R464.

Goring, C. A. I. and W. V. Barthohlomew (1952). "Adsorption of mononucleotides, nucleic acids and nulceoproteins by clays." Soil Science **74**: 149-164.

Greaves, M. and M. Wilson (1969). "The adsorption of nucleic acids by montmorillonite." Soil Biology and Biochemistry **1** (4): 317.
Greaves, M. and M. Wilson (1970). "The degradation of nucleic acids and montmorillonite-nucleic-acid complexes by soil microorganisms." Soil Biology and Biochemistry **2** (4): 257.
Grimaud, P., J. Sauzier, R. Bheekhee and P. Thomas (2006). "Nutritive value of tropical pastures in Mauritius." Tropical Animal Health and Production **38**: 159-167.
Gulden, R. H., S. Lerat, M. M. Hart, J. T. Powell, J. T. Trevors, K. P. Pauls, J. N. Klironomos and C. J. Swanton (2005). "Quantitation of transgenic plant DNA in leachate water: real-time polymerase chain reaction analysis." Journal of Agricultural and Food Chemistry **53**: 5858-5865.
Guo, J., L. X. Du, Y. H. Ma, W. J. Guan, H. B. Li, Q. J. Zhao, X. Li and S. Q. Rao (2005). "A novel maternal lineage revealed in sheep (*Ovis aries*)." Animal Genetics **36** (4): 331-336.
Guthrie, R. (2004). "Radiocarbon evidence of mid-Holocene mammoths stranded on an Alaskan Bering Sea island." Nature: 746-749.
Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, K. L. Han, J. Harshman, C. J. Huddleston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt and T. Yuri (2008). "A phylogenomic study of birds reveals their evolutionary history." Science **320** (5884): 1763-1768.
Haile, J., R. Holdaway, K. Oliver, M. Bunce, M. T. P. Gilbert, R. Nielsen, K. Munch, S. Ho, B. Shapiro and E. Willerslev (2007). "Ancient DNA chronology within sediment deposits: Are paleobiological reconstructions possible and is DNA leaching a factor?" Molecular Biology and Evolution **24** (4): 982-989.
Hall, T. (1999). "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." Nucleic Acids Symposium Series **41**: 95-98.
Handt, O., M. Hoss, M. Krings and S. Paabo (1994). "Ancient DNA: methodological challenges." Experientia **50** (6): 524-9.
Haneca, K., T. Wazny, J. Van Acker and H. Beeckman (2005). "Provenancing Baltic timber from art historical objects: success and limitations." Journal of Archaeological Science **32** (2): 261.
Hansen, A., E. Willerslev, C. Wiuf, T. Mourier and P. Arctander (2001). "Statistical evidence for miscoding lesions in ancient DNA templates." Molecular Biology and Evolution **18** (2): 262-5.
Hansen, A. J., D. L. Mitchell, C. Wiuf, L. Panikert, T. B. Brand, J. Binladen, D. A. Gilichinsky, R. Ronn and E. Willerslev (2006). "Crosslinks rather than strand breaks determine access to ancient DNA sequences from frozen sediments." Genetics **173** (2): 1175-1179.
Hardy, C., C. Callou, J. D. Vigne, D. Casane, N. Dennebouy, J. C. Mounolou and M. Monnerot (1995). "Rabbit mitochondrial DNA diversity from prehistoric to modern times." Journal of Molecular Evolution **40** (3): 227-37.
Harrison, G. L., P. A. McLenachan, M. J. Phillips, K. E. Slack, A. Cooper and D. Penny (2004). "Four new avian mitochondrial genomes help get to basic evolutionary questions in the late cretaceous." Mol Biol Evol **21** (6): 974-83.
Harrison, T. and C. P. Msuya (2005). "Fossil struthionid eggshells from Laetoli, Tanzania: Taxonomic and biostratigraphic significance." Journal of African Earth Sciences **41** (4): 303-315.

Hawes, M. (1990). "Living plant cells released from the root cap: A regulator of microbial populations in the rhizosphere?" Plant and Soil **129** (1): 19-27.
Hawkins, A. F. A. and S. M. Goodman (2003). The Natural History of Madagascar. S. M. Goodman and J. P. Benstead, University of Chicago Press**:** 1019-1044.
Hebsgaard, M. B., M. J. Phillips and E. Willerslev (2005). "Geologically ancient DNA: fact or artefact?" Trends in Microbiology **13** (5): 212-220.
Higham, T. (1994). "Radiocarbon dating New Zealand prehistory with moa eggshell: Some preliminary results." Quaternary Science Reviews **13** (2): 163-169.
Higuchi, R., B. Bowman, M. Freiberger, O. A. Ryder and A. C. Wilson (1984). "DNA sequences from the quagga, an extinct member of the horse family." Nature **312** (5991): 282-4.
Higuchi, R., C. H. von Beroldingen, G. F. Sensabaugh and H. A. Erlich (1988). "DNA typing from single hairs." Nature **332** (6164): 543-6.
Higuchi, R. G., L. A. Wrischnik, E. Oakes, M. George, B. Tong and A. C. Wilson (1987). "Mitochondrial-DNA of the Extinct Quagga - Relatedness and Extent of Postmortem Change." Journal of Molecular Evolution **25** (4): 283-287.
Hofreiter, M., J. Mead, P. Martin and H. Poinar (2003). "Molecular caving." Current Biology **13** (18): R693-5.
Hofreiter, M., H. N. Poinar, W. G. Spaulding, K. Bauer, P. S. Martin, G. Possnert and S. Paabo (2000). "A molecular analysis of ground sloth diet through the last glaciation." Molecular Ecology **9** (12): 1975-84.
Hofreiter, M., D. Serre, H. Poinar, M. Kuch and S. Pääbo (2001). "Ancient DNA." Nature Reviews: Genetics **2**: 353-359.
Holdaway, R. and N. Beavan-Athfield (1999). "Reliable ^{14}C AMS dates on bird and Pacific rat *Rattus exulans* bone gelatin, from a CaCO3-rich deposit." Journal of The Royal Society of New Zealand **29** (3): 185-211.
Holdaway, R. and C. Jacomb (2000). "Rapid extinction of the moas (Aves: Dinornithiformes): model, test, and implications." Science **287** (5461): 2250-4.
Holdaway, R., M. Jones and N. Beavan-Athfield (2002). "Late Holocene extinction of Finsch's duck (*Chenonetta finschi*), an endemic, possibly flightless, New Zealand duck." Journal of The Royal Society of New Zealand **32** (4): 629–651.
Holdaway, R., R. Roberts, N. Beavan-Athfield, J. Olley and T. Worthy (2002). "Optical dating of quartz sediments and accelerator mass spectrometry14C dating of bone gelatin and moa eggshell: a comparison of age estimates for non-archaeological deposits in New Zealand." Journal of The Royal Society of New Zealand **32** (3): 463-505.
Holdaway, R. N. and C. Jacomb (2000). "Rapid extinction of the moas (Aves : Dinorinthiformes): Model, test, and implications." Science **287** (5461): 2250-2254.
Hoss, M., P. Jaruga, T. H. Zastawny, M. Dizdaroglu and S. Paabo (1996). "DNA damage and DNA sequence retrieval from ancient tissues." Nucleic Acids Research **24** (7): 1304-7.
Hubberten, H. W., A. Andreev, V. I. Astakhov, I. Demidov, J. A. Dowdeswell, M. Henriksen, C. Hjort, M. Houmark-Nielsen, M. Jakobsson, S. Kuzmina, E. Larsen, J. P. Lunkka, A. LysÂ, J. Mangerud, P. Mˆller, M. Saarnisto, L. Schirrmeister, A. V. Sher, C. Siegert, M. J. Siegert and J. I. J. I. Svendsen (2004). "The periglacial climate and environment in northern Eurasia during the Last Glaciation." Quaternary Science Reviews **23** (11-13): 1333-1357.

Hulsenbeck, J. and F. Ronquist (2001). "MR BAYES: Bayesian inference of phylogeny." Bioinformatics **17**: 754-755.

Hutton, I. (1990). Birds of Lord Howe Island: Past and Present, Ian Hutton: Coffs Harbour.

Johnson, B. J., G. H. Miller, M. L. Fogel and P. B. Beaumont (1997). "The determination of late Quaternary paleoenvironments at Equus Cave, South Africa, using stable isotopes and amino acid racemization in ostrich eggshell." Palaeogeography, Palaeoclimatology, Palaeoecology **136** (1-4): 121-137.

Kahil Bar-Gal, G., H. Khailaily, O. Mader, P. Ducos and L. Kolska Horwitz (2002). "Ancient DNA evidence for the transition from wild to domestic status in Neolithic goats: a case study from the site of Abu Gosh, Israel." Ancient Biomolecules **4**: 9-17.

Karanth, K. P., T. Delefosse, B. Rakotosamimanana, T. J. Parsons and A. D. Yoder (2005). "Ancient DNA from giant extinct lemurs confirms single origin of Malagasy primates." Proceedings of the National Academy of Sciences of the United States of America **102** (14): 5090-5.

Karanth, K. P. and A. D. Yoder (in press). Geographic origins of ancient DNA samples and prospects for ancient DNA studies in the tropics. Molecular markers, PCR, bioinformatics and ancient DNA—technology, troubleshooting and applications G. Dorado. New York, Science Publishers.

Khanna, M. and G. Stotzky (1992). "Transformation of Bacillus-Subtilis by DNA Bound on Montmorillonite and Effect of Dnase on the Transforming Ability of Bound DNA." Applied and Environmental Microbiology **58** (6): 1930-1939.

Khanna, M., M. Yoder, L. Calamai and G. Stotzky (1998). "X-ray diffractometry and electron microscopy of DNA from *Bacillus subtilis* bound on clay minerals." Sciences of Soils **3** (1).

Kijas, J. W., M. Menzies and A. Ingham (2006). "Sequence diversity and rates of molecular evolution between sheep and cattle genes." Animal Genetics **37** (2): 171-174.

Kornberg, A. (1974). DNA Synthesis, W.H. Freeman and Company, San Francisco.

Krause, D., J. Hartman and N. Wells (1997). Late Cretaceous Vertebrates from Madagascar. Natural Change and Human Impact in Madagascar. S. Goodman and B. Patterson, Smithsonian Institution Press**:** 3-43.

Kuzmin, Y., L. Orlova, I. Zolnikov and A. Igolnikov (2000). "The history of mammoth (Mammuthus primigenius Blum.) population in Siberia and adjacent areas (based on radiocarbon data)." Russian Geology and Geophysics **41** (5): 746-754 (in Russian).

Kwok, S. and R. Higuchi (1989). "Avoiding False Positives with PCR." Nature **339** (6221): 237-238.

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

Larson, G., T. Cucchi, M. Fujita, E. Matisoo-Smith, J. Robins, A. Anderson, B. Rolett, M. Spriggs, G. Dolman, T. H. Kim, N. T. Thuy, E. Randi, M. Doherty, R. A. Due, R. Bollt, T. Djubiantono, B. Griffin, M. Intoh, E. Keane, P. Kirch,

K. T. Li, M. Morwood, L. M. Pedrina, P. J. Piper, R. J. Rabett, P. Shooter, G. Van den Bergh, E. West, S. Wickler, J. Yuan, A. Cooper and K. Dobney (2007). "Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania." Proceedings of the National Academy of Sciences of the United States of America **104** (12): 4834-9.

Larson, G., K. Dobney, U. Albarella, M. Fang, E. Matisoo-Smith, J. Robins, S. Lowden, H. Finlayson, T. Brand, E. Willerslev, P. Rowley-Conwy, L. Andersson and r. C. Coope (2005). "Worldwide Phylogeography of Wild Boar Reveals Multiple Centers of Pig Domestication." Science **307**: 1618-1621.

Leonard, J. A., O. Shanks, M. Hofreiter, E. Kreuz, L. Hodges, W. Ream, R. K. Wayne and R. C. Fleischer (2007). "Animal DNA in PCR reagents plagues ancient DNA research." Journal of Archaeological Science **34** (9): 1361.

Levy-Booth, D. J., R. G. Campbell, R. H. Gulden, M. M. Hart, J. R. Powell, J. N. Klironomos, K. P. Pauls, C. J. Swanton, J. T. Trevors and K. E. Dunfield (2007). "Cycling of extracellular DNA in the soil environment." Soil Biology and Biochemistry **39** (12): 2977-2991.

Lian, O., J. Hu, D. Huntley and S. Hicock (1995). "Optical dating studies of Quaternary organic-rich sediments from southwestern British Columbia and northwestern Washington State." Canadian Journal of Earth Sciences **32**: 1194-1207.

Lian, O. B. and R. G. Roberts (2006). "Dating the Quaternary: progress in luminescence dating of sediments." Quaternary Science Reviews **25** (19-20): 2449.

Lindahl, T. (1977). "DNA N-Glycosidases, a New Class of Enzymes Acting on Damaged DNA." Mutation Research **46** (2): 138-138.

Lindahl, T. (1979). "DNA glycosylases, endonucleases for apurinic/apyrimidinic sitesand base excision-repair." Progress in Nucleic Acid Research and Molecular Biology **22**: 135-192

Lindahl, T. (1993). "Instability and decay of the primary structure of DNA." Nature **362** (6422): 709-15.

Lindahl, T. (1993). "Recovery of Antediluvian DNA." Nature **365** (6448): 700-700.

Lindahl, T., S. Ljungquist, W. Siegert, B. Nyberg and B. Sperens (1977). "DNA N-Glycosidases - Properties of Uracil-DNA Glycosidase from *Escherichia coli*." Journal of Biological Chemistry **252** (10): 3286-3294.

Lindahl, T. and B. Nyberg (1972). "Rate of Depurination of Native Deoxyribonucleic Acid." Biochemistry **11** (19): 3610-&.

Liu, Y. P., G. S. Wu, Y. G. Yao, Y. W. Miao, G. Luikart, M. Baig, A. Beja-Pereira, Z. L. Ding, M. G. Palanichamy and Y. P. Zhang (2006). "Multiple maternal origins of chickens: Out of the Asian jungles." Molecular Phylogenetics and Evolution **38** (1): 12-19.

Long, A., A. Sher and S. Vartanyan (1994). "Holocene mammoth dates." Nature **369**: 364.

Longo, M. C., M. S. Berninger and J. L. Hartley (1990). "Use of Uracil DNA Glycosylase to Control Carry-over Contamination in Polymerase Chain-Reactions." Gene **93** (1): 125-128.

Lorenz, M. and W. Wackernagel (1994). "Bacterial gene transfer by natural genetic transformation in the environment." Microbiological Reviews **58** (3): 563-602.

Lorenz, M. G. and W. Wackernagel (1987). "Adsorption of DNA to Sand and Variable Degradation Rates of Adsorbed DNA." Applied and Environmental Microbiology **53** (12): 2948-2952.

Lorenz, M. G. and W. Wackernagel (1991). "High Frequency of Natural Genetic Transformation of *Pseudomonas stutzeri* in Soil Extract Supplemented with a Carbon/Energy and Phosphorus Source." Applied and Environmental Microbiology **57** (4): 1246-1251.

Lydolph, M., J. Jacobsen, P. Arctander, M. Gilbert, D. Gilichinsky, A. Hansen, E. Willerslev and L. Lange (2005). "Beringian paleoecology inferred from permafrost-preserved fungal DNA." Applied and Environmental Microbiology **71** (2): 1012-7.

MacPhee, R. and P. Marx (1997). The 40,000-Year Plague: Humans, Hyperdisease, and First-Contact Extinctions. Natural Change and Human Impact in Madagascar. S. M. Goodman and B. Patterson, Smithsonian Institution Press.

MacPhee, R. D. E., A. N. Tikhonov, D. Mol, C. D. Maliave, H. Van der Plicht, A. D. Greenwood, C. Flemming and L. Agenbroad (2002). "Radiocarbon chronologies and extinction dynamics of the Late Quaternary mammalian megafauna of the Taimyr Peninsula, Russian Federation." Journal of Archaeological Science **29** (9): 1017-1042.

Mahé, J. and M. Sourdat (1972). "Sur l'extinction des vertébrés subfossiles et l'aridification du climat dans le sud-ouest de Madagascar." Bulletin de la Société Géologique de France **14**: 295-309.

Martin, P. (1984). Prehistoric Overkill: The Global Model, University of Arizon Press, Tuscon.

Martin, P. (2005). Twilight of the Mammoths. Berkeley, University of California Press.

Matisoo-Smith, E., R. M. Roberts, G. J. Irwin, J. S. Allen, D. Penny and D. M. Lambert (1998). "Patterns of prehistoric human mobility in polynesia indicated by mtDNA from the Pacific rat." Proceedings of the National Academy of Sciences of the United States of America **95** (25): 15145-50.

Mawer, G. (1999). Ahab's Trade: the saga of south seas whaling, St Leonards: Allen and Unwin.

McGeehin, J., G. S. Burr, A. J. T. Jull, D. Reines, J. Gosse, P. T. Davis, D. Muhs and J. R. Southon (2001). "Stepped-combustion C-14 dating of sediment: A comparison with established techniques." Radiocarbon **43** (2A): 255-261.

Meier, P. and W. Wackernagel (2003). "Monitoring the spread of recombinant DNA from field plots with transgenic sugar beet plants by PCR and natural transformation of Pseudomonas stutzeri." Transgenic Research **12** (3): 293-304.

Miller, G. H., P. B. Beaumont, H. J. Deacon, A. S. Brooks, P. E. Hare and A. J. T. Jull (1999). "Earliest modern humans in southern Africa dated by isoleucine epimerization in ostrich eggshell." Quaternary Science Reviews **18** (13): 1537-1548.

Mishra, N. (2002). Nucleases: Molecular biology and applications, Wiley, Hoboken, NJ**:** 45-60.

Mitchell, D., E. Willerslev and A. Hansen (2005). "Damage and repair of ancient DNA." Mutation Research **571** (1-2): 265-276.

Moller, P., D. Y. Bolshiyanov and H. Bergsten (1999). "Weichselian geology and palaeoenvironmental history of the central Taymyr Peninsula, Siberia, indicating no glaciation during the last global glacial maximum." Boreas **28** (1): 92-114.

Monnier, L. and C. Lamberton (1955). "Note sur des ossements sub-fossiles de la région de Manajary." Bulletin de l'Académie Malgache **nouvelle série 3**: 211-13.
Moore, P. D., J. A. Webb and M. E. Collinson (1991). Pollen analysis, Blackwell Scientific Publications.
Morat, P. (1973). "Les Savanes de sud-ouest de Madagascar." Memoires O.R.S.T.O.M. **no. 68**.
Mullis, K. and F. Faloona (1987). "Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction." " Methods in Enzymology **155**: 335-50.
Newman, M. E., J. S. Parboosingh, P. J. Bridge and H. Ceri (2002). "Identification of Archaeological Animal Bone by PCR/DNA Analysis." Journal of Archaeological Science **29** (1): 77.
Nicholls, M. (1952). A History of Lord Howe Island. Hobart, Mercury Press.
Nichols, D. (2006). Lord Howe Island Rising. Frenchs Forest, Tower Books.
Nielsen, K. M., A. M. Bones and J. D. Van Elsas (1997). "Induced Natural Transformation of Acinetobacter calcoaceticus in Soil Microcosms." Applied and Environmental Microbiology **63** (10): 3972-3977.
Nielsen, K. M., M. D. van Weerelt, T. N. Berg, A. M. Bones, A. N. Hagler and J. D. van Elsas (1997). "Natural transformation and availability of transforming DNA to Acinetobacter calcoaceticus in soil microcosms." Applied and Environmental Microbiology **63** (5): 1945-52.
Noonan, J. P., M. Hofreiter, D. Smith, J. R. Priest, N. Rohland, G. Rabeder, J. Krause, J. C. Detter, S. Paabo and E. M. Rubin (2005). "Genomic sequencing of Pleistocene cave bears." Science **309** (5734): 597-600.
Noro, M., R. Masuda, I. A. Dubrovo, M. C. Yoshida and M. Kato (1998). "Molecular phylogenetic inference of the woolly mammoth Mammuthus primigenius, based on complete sequences of mitochondrial cytochrome b and 12S ribosomal RNA genes." Journal of Molecular Evolution **46** (3): 314-326.
Ogram, A., G. Sayler, D. Gustin and R. Lewis (1988). "DNA adsorption to soils and sediments." Environmental Science and Technology **22**: 982 - 984.
Olley, J. M., T. Pietsch and R. G. Roberts (2004). "Optical dating of Holocene sediments from a variety of geomorphic settings using single grains of quartz." Geomorphology **60** (3-4): 337-358.
Otto-Bliesner, B., E. Brady, G. Clauzet, R. Tomas, S. Levis and Z. Kothavala (2006). "Last Glacial Maximum and Holocene Climate in CCSM3." Journal of Climate **19**: 2526-2544.
Pääbo, S., R. G. Higuchi and A. C. Wilson (1989). "Ancient DNA and the Polymerase Chain-Reaction - the Emerging Field of Molecular Archaeology." Journal of Biological Chemistry **264** (17): 9709-9712.
Pääbo, S., D. M. Irwin and A. C. Wilson (1990). "DNA Damage Promotes Jumping between Templates During Enzymatic Amplification." Journal of Biological Chemistry **265** (8): 4718-4721.
Pääbo, S., H. Poinar, D. Serre, V. Jaenicke-Despres, J. Hebler, N. Rohland, M. Kuch, J. Krause, L. Vigilant and M. Hofreiter (2004). "Genetic Analyses from Ancient DNA." Annual Review of Genetics **38**: 645-679.
Padya, B. M. (1989). The Climate of Mauritius, Meteorological Office, Mauritius.
Palmen, R. and K. J. Hellingwerf (1997). "Uptake and processing of DNA by *Acinetobacter calcoaceticus:* a review." Gene **192** (1): 179-90.
Parducci, L. and R. J. Petit (2004). "Ancient DNA - unlocking plants' fossil secrets." New Phytologist **161** (2): 335-339.

Pastorini, J., U. Thalmann and R. D. Martin (2003). "A molecular approach to comparative phylogeography of extant Malagasy lemurs." Proceedings of the National Academy of Sciences of the United States of America **100** (10): 5879-5884.
Phillips, M. J., G. Gibb, E. Crimp and D. Penny (in prep). "Tinamou and Moa Flock Together: Mitochondrial Genome Sequence Analysis Reveals Independent Losses of Flight Among Ratites."
Pietramellara, G., M. T. Ceccherini, J. Ascher and P. Nannipieri (2006). "Persistence of transgenic and not transgenic extracellular DNA in soil and bacterial transformation." Rivista di Biologia **99** (1): 37-68.
Pietramellara, G., M. Franchi, E. Gallori and P. Nannipieri (2001). "Effect of molecular characteristics of DNA on its adsorption and binding on homoionic montmorillonite and kaolinite." Biology and Fertility of Soils **33** (5): 402-409.
Pischetsrieder, M., W. Seidel, G. Munch and R. Schinzel (1999). "N-2-(1-carboxyethyl)deoxyguanosine, a nonenzymatic glycation adduct of DNA, induces single-strand breaks and increases mutation frequencies." Biochemical and Biophysical Research Communications **264** (2): 544-549.
Poinar, H. (2003). "The Top 10 List: criteria of authenticity for DNA from ancient and forensic samples." International Congress Series **1239**: 575-579.
Poinar, H., R. Cano and G. Poinar (1993). "DNA from an extinct plant " Nature **363**: 677.
Poinar, H., M. Hoss, J. Bada and S. Paabo (1996). "Amino acid racemization and the preservation of ancient DNA." Science **272** (5263): 864-6.
Poinar, H. N., M. Hofreiter, W. G. Spaulding, P. S. Martin, B. A. Stankiewicz, H. Bland, R. P. Evershed, G. Possnert and S. Paabo (1998). "Molecular coproscopy: dung and diet of the extinct ground sloth *Nothrotheriops shastensis*." Science **281** (5375): 402-6.
Polverari, A., R. Buonaurio, S. Guiderdone, M. Pezzotti and M. Marte (2000). "Ultrastructural observations and DNA degradation analysis of pepper leaves undergoing a hypersensitive reaction to *Xanthomonas campestris pv. vesicatoria*." European Journal of Plant Pathology **106** (5): 423-431.
Pompanon, F., E. Pettex and L. Desprès (2006). "Patterns of resource exploitation in four coexisting globeflower fly species (*Chiastocheta sp.*)." Acta Oecologica **29** (2): 233-240.
Poté, J., M. T. Ceccherini, V. T. Van, W. Rosselli, W. Wildi, P. Simonet and T. M. Vogel (2003). "Fate and transport of antibiotic resistance genes in saturated soil columns." European Journal of Soil Biology **39** (2): 65-71.
Poté, J., P. Rosse, W. Rosselli, V. T. Van and W. Wildi (2005). "Kinetics of mass and DNA decomposition in tomato leaves." Chemosphere **61** (5): 677-684.
Poulakakis, N., A. Parmakelis, P. Lymberakis, M. Mylonas, E. Zouros, D. S. Reese, S. Glaberman and A. Caccone (2006). "Ancient DNA forces reconsideration of evolutionary history of Mediterranean pygmy elephantids." Biology Letters **2** (3): 451-4.
Prentice, C. (1980). "Multidimensional scaling as a research tool in Quaternary palynology: a review of theory and methods." Review of Palaeobotany and Palynology **31**: 71–104.
Quammen, D. (1996). The Song of the Dodo, Hutchinson.
Rabone, H. (1940). Lord Howe Island: its discovery and early associations 1788 to 1888. Sydney, Australis.

Ray, N. and J. Adams (2001) "A GIS-based Vegetation Map of the World at the Last Glacial Maximum (25,000-15,000 BP)." Internet Archaeology **11**, DOI:
Rijsdijk, K. F., J. P. Hume, F. Bunnik, F. B. Vincent Florens, C. Baider, B. Shapiro, H. van der Plicht, A. Janoo, O. Griffiths, L. W. van den Hoek Ostende, H. Cremer, J. Haile, T. Vernimmen and E. Gittenberger ((in prep.)). "Reconstructing the world of the dodo." Quaternary Science Reviews.
Robinson, D. and N. King (1968). "Mucopolsaccharides of an avian egg shell membrane." Journal of the Royal Microscopical Society **88**: 13-22.
Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis and H. A. Erlich (1988). "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase." Science **239** (4839): 487-91.
Salamon, M., N. Tuross, B. Arensburg and S. Weiner (2005). "Relatively well preserved DNA is present in the crystal aggregates of fossil bones." Proceedings of the National Academy of Sciences of the United States of America **102** (39): 13783-13788.
Sampson, C. G. (1994). "Ostrich eggs and Bushman survival on the north-east frontier of the Cape Colony, South Africa." Journal of Arid Environments **26** (4): 383-399.
Savolainen, P., Y. P. Zhang, J. Luo, J. Lundeberg and T. Leitner (2002). "Genetic evidence for an East Asian origin of domestic dogs." Science **298** (5598): 1610-1613.
Schweitzer, M. H., Z. Suo, R. Avci, J. M. Asara, M. A. Allen, F. T. Arce and J. R. Horner (2007). "Analyses of soft tissue from Tyrannosaurus rex suggest the presence of protein." Science **316** (5822): 277-280.
Shapiro, B., A. Drummond, A. Rambaut, M. Wilson, P. Matheus, A. Sher, O. Pybus, M. Gilbert, I. Barnes, J. Binladen, E. Willerslev, A. Hansen, G. Baryshnikov, J. Burns, S. Davydov, D. J, D. Froese, C. Harington, G. Keddie, P. Kosintsev, P. Kunz, L. Martin, R. Stephenson, R. Storer, R. Tedford, S. Zimov and A. Cooper (2004). "Rise and Fall of the Beringian Steppe Bison." Science **306**: 1561-1565.
Smith, A., D. Smith and B. Funnell (1994). Atlas of Mesozoic and Cenozoic coastlines, Cambridge University Press.
Smith, C. I., A. T. Chamberlain, M. S. Riley, C. Stringer and M. J. Collins (2003). "The thermal history of human fossils and the likelihood of successful DNA amplification." Journal of Human Evolution **45** (3): 203-17.
Sogin, M. L., H. G. Morrison, J. A. Huber, D. M. Welch, S. M. Huse, P. R. Neal, J. M. Arrieta and G. J. Herndl (2006). "Microbial diversity in the deep sea and the underexplored "rare biosphere"." Proceedings of the National Academy of Sciences of The United States of America **103** (32): 12115-12120.
Solow, A. R., D. L. Roberts and K. M. Robbirt (2006). "On the Pleistocene extinctions of Alaskan mammoths and horses." Proceedings of the National Academy of Sciences of the United States of America **103** (19): 7351-7353.
Sønstebø, J., L. Gielly, A. Brysting, R. Elven, M. Edwards, J. Haile, E. Willerslev, K. Bråthen, N. Yoccoz, E. Coissac, D. Rioux, J. Sannier, P. Taberlet and C. Brochmann (in prep.). "A minimalist DNA barcoding approach for reconstructing past arctic vegetation and climate."
Sparks, R., W. Melhuish, J. McKee, J. Ogden, J. Palmer and B. Molloy (1995). "^{14}C calibration in the southern hemisphere and the date of the last Taupo eruption: evidence from tree-ring sequences." Radiocarbon **37**: 155-163.

Staub, F. (1993). Fauna of Mauritius and associated flora. Port Louis, Précigraph Limited.
Stidham, T. A. and J. K. Brophy (2007). "Relative dating of South African Middle Stone Age (MSA) sites using variation in ostrich eggshell morphology." American Journal of Physical Anthropology: 225-226.
Stoddardt, D. R., J. F. Peake, C. Gordon and R. Burleigh (1979). "Historical records of Indian Ocean giant tortoise populations." Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **286**: 147-158.
Stotzky, G., M. A. Devanas and L. R. Zeph (1989). "Behavior of genetically engineered microbes in natural environments and their potential use in in situ reclamation of contaminated sites." Schriftenr Ver Wasser Boden Lufthyg **80**: 293-343.
Strausberger, B. M. and M. V. Ashley (2001). "Eggs yield nuclear DNA from egg-laying female cowbirds, their embryos and offspring." Conservation Genetics **2** (4): 385-390.
Sulerzhitsky, L. (1997). Patterns of the radiocarbon chronology of mammoths in Siberia and northern Eastern Europe (as sub-stratum for human dispersal), Institute of Geography, Moscow.
Sulerzhitsky, L. and F. Romanenko (1997). "Age and distribution of the "mammoth" fauna of the polar regions of Asia." Earth Cryosphere **1** (4): 12-19 (in Russian).
Swofford, D. L. (1993). "Paup - a Computer-Program for Phylogenetic Inference Using Maximum Parsimony." Journal of General Physiology **102** (6): A9-A9.
Taberlet, P. and J. Bouvet (1991). "A Single Plucked Feather as a Source of DNA for Bird Genetic-Studies." Auk **108** (4): 959-960.
Taberlet, P., E. Coissac, F. Pompanon, L. Gielly, C. Miquel, A. Valentini, T. Vermat, G. Corthier, C. Brochmann and E. Willerslev (2007). "Power and limitations of the chloroplast trnL(UAA) intron for plant DNA barcoding." Nucleic Acids Research **35** (3).
Taberlet, P., H. Mattock, C. Dubois-Paganon and J. Bouvet (1993). "Sexing free-ranging brown bears Ursus arctos using hairs found in the field." Molecular Ecology **2** (6): 399-403.
Taylor, P. G. (1996). "Reproducibility of ancient DNA sequences from extinct Pleistocene fauna." Molecular Biology and Evolution **13** (1): 283-285.
Thomas, C. and K. Nielsen (2005). "Mechanisms of, and Barriers to, Horizontal Gene Transfer between Bacteria." Nature Reviews Microbiology **3**: 711-721.
Trevors, J. T. (1996). "DNA in soil: Adsorption, genetic transformation, molecular evolution and genetic microchip." Antonie Van Leeuwenhoek International Journal Of General And Molecular Microbiology **70** (1): 1-10.
Troy, C. S., D. E. MacHugh, J. F. Bailey, D. A. Magee, R. T. Loftus, P. Cunningham, A. T. Chamberlain, B. C. Sykes and D. G. Bradley (2001). "Genetic evidence for Near-Eastern origins of European cattle." Nature **410** (6832): 1088-91.
Turbin, N. B. (1965). "Nucleic acid content in pollen of diloid and tetrapoloid sugar beet." Doklady Akademii Nauk SSSR **161**: 463-465.
Turner, B. L. and S. Newman (2005). "Phosphorus cycling in wetland soils: the importance of phosphate diesters." Journal of Environmetal Quality **34** (5): 1921-9.
Vadehra, D., R. Baker and H. Naylor (1972). "Distribution of lysozyme activity in the exteriors of eggs from *Gallus gallus*." Comparative Biochemistry and Physiology **43B**: 503-508.

Valiere, N. and P. Taberlet (2000). "Urine collected in the field as a source of DNA for species and individual identification." Molecular Ecology **9** (12): 2150-2.
van der Knaap, W. O., J. F. N. van Leeuwen, W. Finsinger, E. Gobet, R. Pini, A. Schweizer, V. Valsecchi and B. Ammann (2005). "Migration and population expansion of *Abies, Fagus, Picea,* and *Quercus* since 15000 years in and across the Alps, based on pollen-percentage threshold values." Quaternary Science Reviews **24** (5-6): 645-680.
Vartanyan, S., K. Arslanov, T. Tertychynaya and S. Chernov (1995). "Radiocarbon evidence for mammoths on Wrangel Island Arctic Ocean until 2000BC." Radiocarbon **37** (1): 7-10.
Vartanyan, S., V. Garrut and A. Sher (1993). "Holocene dwarf mammoths from Wrangel Island in the Siberian Arctic." Nature **382**: 337-340.
Vasan, S., X. Zhang, X. N. Zhang, A. Kapurniotu, J. Bernhagen, S. Teichberg, J. Basgen, D. Wagle, D. Shih, I. Terlecky, R. Bucala, A. Cerami, J. Egan and P. Ulrich (1996). "An agent cleaving glucose-derived protein crosslinks in vitro and in vivo." Nature **382** (6588): 275-278.
Veltre, D. W. e. a. (in press). Quaternary Research.
Vila, C., J. A. Leonard, A. Gotherstrom, S. Marklund, K. Sandberg, K. Liden, R. K. Wayne and H. Ellegren (2001). "Widespread origins of domestic horse lineages." Science **291** (5503): 474-7.
Voet, D. and J. Voet (1995). Biochemistry, John Wiley & Sons: New York.
Vogel, J. C., E. Visser and A. Fuls (2001). "Suitability of ostrich eggshell for radiocarbon dating." Radiocarbon **43** (1): 133-137.
von der Heyden, S. and T. Cavalier-Smith (2005). "Culturing and environmental DNA sequencing uncover hidden kinetoplastid biodiversity and a major marine clade within ancestrally freshwater *Neobodo designis*." International Journal of Systematic and Evolutionary Microbiology **55**: 2605-2621.
Walker, D. A., M. K. Raynolds, F. J. A. Daniels, E. Einarsson, A. Elvebakk, W. A. Gould, A. E. Katenin, S. S. Kholod, C. J. Markon, E. S. Melnikov, N. G. Moskalenko, S. S. Talbot, B. A. Yurtsev and C. Team (2005). "The Circumpolar Arctic vegetation map." Journal of Vegetation Science **16** (3): 267-282.
Wang, H., J. Li, R. M. Bostock and D. G. Gilchrist (1996). "Apoptosis: A functional paradigm for programmed plant cell death induced by a host-selective phytotoxin and invoked during development." Plant Cell **8** (3): 375-391.
Wayne, R. K., J. A. Leonard and A. Cooper (1999). "FULL OF SOUND AND FURY: History of Ancient DNA." Annual Review of Ecology and Systematics **30**: 457-477.
Wetmore, A. (1967). "Re-Creating Madagascars Giant Extinct Bird." National Geographic **132** (4): 488-493.
Wheeler, D., C. Chappey, A. Lash, D. Leipe, T. Madden, G. Schuler, T. Tatusova and B. Rapp (2000). "Database resources of the National Center for Biotechnology Information." Nucleic Acids Research **28** (1): 10-4.
Widmer, F., R. J. Seidler, K. K. Donegan and G. L. Reed (1997). "Quantification of transgenic plant marker gene persistence in the field." Molecular Ecology **6** (1): 1-7.
Willerslev, E., E. Cappellini, W. Boomsma, R. Nielsen, M. B. Hebsgaard, T. B. Brand, M. Hofreiter, M. Bunce, H. N. Poinar, D. Dahl-Jensen, S. Johnsen, J. P. Steffensen, O. Bennike, J. L. Schwenninger, R. Nathan, S. Armitage, C. J. de Hoog, V. Alfimov, M. Christl, J. Beer, R. Muscheler, J. Barker, M. Sharp,

K. E. Penkman, J. Haile, P. Taberlet, M. T. Gilbert, A. Casoli, E. Campani and M. J. Collins (2007). "Ancient biomolecules from deep ice cores reveal a forested southern Greenland." Science **317** (5834): 111-4.

Willerslev, E. and A. Cooper (2005). "Ancient DNA." Proc Biol Sci **272** (1558): 3-16.

Willerslev, E., A. Hansen, J. Binladen, T. Brand, M. Gilbert, B. Shapiro, M. Bunce, C. Wiuf, D. Gilichinsky and A. Cooper (2003). "Diverse plant and animal genetic records from Holocene and Pleistocene sediments." Science **300** (5620): 791-5.

Willerslev, E., A. Hansen, B. Christensen, J. Steffensen and P. Arctander (1999). "Diversity of Holocene life forms in fossil glacier ice." Proceedings of the National Academy of Sciences of the United States of America **96** (14): 8017-21.

Willerslev, E., A. Hansen and H. Poinar (2004). "Isolation of nucleic acids and cultures from fossil ice and permafrost." Trends in Ecology and Evolution **19** (3): 141-147

Willerslev, E., A. Hansen, R. Rønn, T. Brand, I. Barnes, C. Wiuf, D. Gilichinsky, D. Mitchell and A. Cooper (2004). "Long-term persistence of bacterial DNA." Current Biology **14** (1): R9-R10.

Winn, S. and H. Ball (1975). "β-N-Acetylglucosaminidase activity of the albumen layers and membranes of the chicken's egg." Poultry Science **54**: 799-805.

Wintle, A. G. and A. S. Murray (2006). "A review of quartz optically stimulated luminescence characteristics and their relevance in single-aliquot regeneration dating protocols." Radiation Measurements **41** (4): 369-391.

Woodward, S. R., N. J. Weyand and M. Bunnell (1994). "DNA-Sequence from Cretaceous Period Bone Fragments." Science **266** (5188): 1229-1232.

Worthy, T. (1990). "An analysis of the distribution and relative abundance of moa species (Aves: Dinornithiformes)." New Zealand Journal of Zoology **17**: 213-241.

Worthy, T. and R. Holdaway (1993). "Quaternary fossil faunas from caves in the Punakaiki area, West Coast, South Island, New Zealand." Journal of the Royal Society of New Zealand **23**: 147-254.

Worthy, T. and R. Holdaway (2000). "Terrestrial fossil vertebrate faunas from inland Hawke's Bay, North Island, New Zealand." *Records of the Canterbury Museum **14***: 89-154.

Worthy, T. and S. Swabey (2002). "Avifaunal changes revealed in Quaternary deposits near Waitomo Caves, North Island, New Zealand." Journal of the Royal Society of New Zealand **32** (2): 293-325.

Zazula, G. D., A. M. Telka, C. R. Harington, C. E. Schweger and R. W. Mathewes (2006). "New spruce (*Picea spp.*) macrofossils from Yukon Territory: Implications for Late Pleistocene refugia in Eastern Beringia." Arctic **59** (4): 391-400.

Appendices

Appendix 1. Showing species codes used in Table 4.2

Species	Code
Agrostis mertensii	*Agr mer*
Agrostis	*Agr sp*
Alchemilla alpina	*Alc alp*
Alchemilla sp.	*Alc sp*
Alnus incana	*Aln inc*
Andromeda polifolia	*And pol*
Angelica archangelica ssp. archangelica	*Ang arc arc*
Anthoxanthum odoratum ssp. alpinum	*Ant odo alp*
Anthriscus sylvestris	*Ant syl*
Arctostaphylos alpinus	*Arc alp*
Athyrium distentifolium	*Ath dis*
Bartsia alpina	*Bar alp*
Betula nana	*Bet nan*
Betula pubescens coll.	*Bet pub coll*
Bistorta vivipara	*Bis viv*
Botrychium lunaria	*Bot lun*
Calamagrostis lapponica	*Cal lap*
Calamagrostis purpurea	*Cal pur*
Calamatrostis	*Cal sp*
Calamagrostis stricta	*Cal str*
Calamagrostis stricta/lapponica	*Cal stri lap*
Caltha palustris coll.	*Cal pal coll*
Campanula rotundifolia coll.	*Camp rot coll*
Carex aquatilis x bigelowii	*Car aqu big*
Carex aquatilis coll.	*Car aqu coll*
Carex aquatilis x juncella	*Car aqu jun*
Carex aquatilis ssp. stans	*Car aqu sta*
Carex bigelowii	*Car big*
Carex brunnescens	*Car bru*
Carex brunnescens/cannescens	*Car bru can*
Carex canescens	*Car can*
Carex lachenalii	*Car lac*
Carex limosa	*Car lim*
Carex limosa x rariflora	*Car lim x rar*
Carex livida	*Car liv*
Carex nigra ssp. juncella	*Car nig jun*
Carex nigra ssp. nigra	*Car nig nig*
Cardamine palustris	*Car pra coll*
Carex saxatilis	*Car sax*
Carex sp	*Car sp*
Carex vaginata	*Car vag*
Cassiope hypnoides	*Cas hyp*
Cerastium alpinum coll.	*Cer alp coll*
Cerastium cerastoides	*Cer cer*
Cerastium fontanum coll	*Cer fon coll*
Cerastium sp	*Cer sp*
Cirsium helenioides	*Cir hel*
Cornus suecica	*Cor suec*
Deschampsia alpina	*Des alp*

Species	Code
Deschampsia cespitosa	*Des ces*
Deschampsia flexuosa	*Des fle*
Empetrum nigrum ssp. hermaphroditum	*Emp nig her*
Epilobium anagallidifolium	*Epi ana*
Epilobium angustifolium	*Epi ang*
Epilobium hornemannii	*Epi hor*
Epilobium palustre	*Epi pal*
Epilobium sp.	*Epi sp*
Equisetum arvense	*Equ arv*
Equisetum pratense	*Equ pra*
Equisetum sp.	*Equ sp*
Equisetum sylvaticum	*Equ syl*
Equisetum variegatum	*Equ var*
Eriophorum angustifolium	*Eri ang*
Euphrasia frigida	*Eup fri*
Euphrasia sp	*Eup sp*
Festuca ovina	*Fes ovi*
Festuca rubra coll.	*Fes rub coll*
Festuca sp.	*Fes sp*
Geranium sylvaticum	*Ger syl*
Geum rivale	*Geu riv*
Hieracium seksjon Alpina	*Hie Alp*
Hieracium sp.	*Hie sp*
Hieracium seksjon Sylvatica	*Hie Syl*
Huperzia selago	*Hup sel*
Juncus filiformis	*Jun fil*
Juncus trifidus	*Jun tri*
Juniperus communis	*Juni com*
Linnaea borealis	*Lin bor*
Loiseleuria procumbens	*Loi pro*
Luzula multiflora coll.	*Luz mul*
Luzula sp.	*Luz sp*
Luzula sudetica	*Luz sud*
Lychnis alpina	*Lyc ann*
Lycopodium clavatum coll.	*Lyc cla coll*
Melampyrum pratense	*Mel pra*
Melampyrum sp.	*Mel sp*
Melampyrum sylvaticum	*Mel syl*
Milium effusum	*Mil eff*
Minuartia	*Min sp*
Myosotis	*Myo sp*
Nardus stricta	*Nar stri*
Omalotheca norvegica	*Oma nor*
Omalotheca supina	*Oma sup*
Oxyria digyna	*Oxy dig*
Parnassia palustris coll.	*Par pal coll*
Pedicularis lapponica	*Ped lap*
Pedicularis	*Ped sp*

Species	Code
Phleum alpinum	*Phl alp*
Phyllodoce caerulea	*Phy cae*
Poa alpina coll.	*Poa alp coll*
Poa arctica coll.	*Poa arc coll*
Poa pratensis ssp. alpigena	*Poa pra alp*
Poa	*Poa sp*
Potentilla palustris	*Pot pal*
Pyrola minor	*Pyr min*
Pyrola sp.	*Pyr sp*
Ranunculus acris coll.	*Ran acr coll*
Ranunculus	*Ran sp*
Rhinanthus minor coll.	*Rhi min coll*
Rhodiola rosea	*Rho ros*
Rubus chamaemorus	*Rub cha*
Rumex acetosa coll.	*Rum ace coll*
Sagina saginoides	*Sag sag*
Sagina sp.	*Sag sp*
Salix glauca coll.	*Sal gla coll*
Salix glauca x phylicifolia	*Sal gla phy*
Salix hastata x lanata	*Sal has lan*
Salix herbacea	*Sal her*
Salix lanata coll.	*Sal lan*
Salix lapponum	*Sal lap*
Salix phylicifolia	*Sal phy*
Salix	*Sal sp*
Saussurea alpina	*Sau alp*
Selaginella selaginoides	*Sel sel*
Sibbaldia procumbens	*Sib pro*
Solidago virgaurea	*Sol vir*
Stellaria borealis	*Ste bor*
Stellaria media	*Ste med*
Stellaria nemorum	*Ste nem*
Stellaria sp.	*Ste sp*
Taraxacum croceum-gruppa	*Tar cro*
Taraxacum sp.	*Tar sp*
Thalictrum alpinum	*Tha alp*
Trientalis europaea	*Tri eur*
Trisetum spicatum	*Tri spi*
Trollius europaeus	*Tro eur*
Vaccinium myrtillus	*Vac myr*
Vaccinium uliginosum	*Vac uli*
Vaccinium vitis-idaea	*Vac vit*
Vahlodea atropurpurea	*Vah atr*
Veronica alpina ssp. alpina	*Ver alp alp*
Veronica	*Ver sp*
Viola biflora	*Vio bif*
Viola palustris	*Vio pal*
Viola sp.	*Vio sp*

Appendix 2. Plant species identified from DNA extracts of Modern Sediments

ENG A 1.1

Ligation	Fusion	Family	Genus	Species
	49			*Avenella flexuosa*
90	47	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
17	23			*Bistorta vivipara*
	19		*Empetrum*	*sibiricum / subholarcticum*
	18		*Hippuris*	*vulgaris / tetraphylla*
53	14		*Rumex*	5
	10			*Thymus oxyodontus*
	10			*Vaccinium alaskense*
	9			*Phleum alpinum*
16	6		*Deschampsia*	5
	6			*Sorbus aucuparia*
	4	Rosaceae	2	*Rosa acicularis / Rubus sp.*
	4		*Huperzia*	*arctica / haleakalae*
8	4		*Ranunculus*	*acris / subborealis / turneri*
	4			*Anthoxanthum nipponicum*
	4			*Trientalis europaea*
	3	Apiaceae	4	Apiaceae
	3			*Alchemilla glomerulans*
	3			*Trollius europaeus*
	2	Asteraceae	5	Asteraceae 3
	2	Asteraceae	4	Asteraceae 2
28	2		*Calamagrostis*	7
	2		*Dryas*	6
	2		*Picea*	*glauca / mariana*
	2		*Vaccinium*	*alaskense / myrtillus*
	2			*Hulteniella integrifolia*
	2			*Noccaea arctica*
	2			*Ranunculus arcticus*
	2			*Vaccinium uliginosum*
	2			*Viola biflora*
	1	Asteraceae	3	Asteraceae 4
	1	Poaceae	8	Poaceae 2
	1			*Cornus suecica*

ENG A 1.2

Ligation	Fusion	Family	Genus	Species
92	437			*Viola biflora*
508	177		*Deschampsia*	5
181	64			*Avenella flexuosa*
11	16		*Ranunculus*	*acris / subborealis / turneri*
17	5		*Rumex*	5
2	4		*Festuca*	9
8	2			*Bistorta vivipara*
26	2	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
5	2	Poaceae	8	Poaceae 2
9	1		*Calamagrostis*	7
5		Asteraceae		
1				*Botrychium lunaria*

ENG A 2.2

Ligation	Fusion	Family	Genus	Species
454	326		*Empetrum*	*sibiricum / subholarcticum*
	30			*Vaccinium alaskense*
59	6		*Vaccinium*	*alaskense / myrtillus*
	2		*Agrostis*	
	1	Asteraceae	3	Asteraceae 4
	1		*Juniperus*	*communis / horizontalis*
	1			*Avenella flexuosa*
1	1			*Vaccinium myrtillus*

ENG A 2.3

Ligation	Fusion	Family	Genus	Species
38	420			*Viola biflora*
1128	392			*Avenella flexuosa*
150	42	Asteraceae	3	Asteraceae 4
32	10	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
56	8		*Calamagrostis*	7
23	8		*Festuca*	9
143	5		*Deschampsia*	5
14	4			*Bistorta vivipara*
	1			*Omalotheca norvegica*
	1			*Phleum alpinum*
part of 7	1	Poaceae	2	*Elymus* sp. / *Leymus* sp.
part of 7	1	Poaceae	8	Poaceae 2

ENG A 3.1

Ligation	Fusion	Family	Genus	Species
68	491			*Viola biflora*
227	131			*Avenella flexuosa*
132	58		*Calamagrostis*	7
21	37		*Rumex*	5
30	29	Poaceae	8	Poaceae 2
28	20		*Deschampsia*	9
33	16	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
11	8			*Phleum alpinum*
	2			*Beckmannia syzigachne*
28	1		*Festuca*	5
3			*Viola*	
1				*Poa alpina*

ENG_A_1.3

Ligation	Fusion	Family	Genus	Species
1342	629			*Avenella flexuosa*
138	177		*Ranunculus*	*acris / subborealis / turneri*
248	59			*Anthoxanthum nipponicum*
83	14		*Calamagrostis*	7
	11	Lycopodiaceae	2	*Diphasiastrum alpinum / Lycopodium sp.*
19	6			*Poa alpina*
3	6	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
15	5		*Festuca*	9
	4	Poaceae	2	*Elymus* sp. / *Leymus* sp.
20	4		*Rumex*	5
	3	Asteraceae	3	Asteraceae 4
	2			*Bistorta vivipara*
	2	Poaceae	8	Poaceae 2
	1			*Pedicularis oederi*

ENG_A_2.1

Ligation	Fusion	Family	Genus	Species
497	169		*Empetrum*	*sibiricum / subholarcticum*
	64			*Vaccinium alaskense*
4	18			*Vaccinium myrtillus*
44	8		*Vaccinium*	*alaskense / myrtillus*
	2			*Vaccinium vitisidaea*
	1	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
2				*Sambucus racemosa*
1				*Cornus suecica*

ENG_B_1.1

Ligation	Fusion	Family	Genus	Species
169	511			*Bistorta vivipara*
180	326			*Avenella flexuosa*
125	289	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
112	136	Asteraceae	4	Asteraceae 2
6	69		*Ranunculus*	*acris / subborealis / turneri*
88	58			*Alchemilla glomerulans*
68	39			*Anthoxanthum nipponicum*
24	30		*Festuca*	9
	22		*Equisetum*	*arvense / fluviatile / sylvaticum*
4	22		*Deschampsia*	5
1	13			*Viola biflora*
	9			*Poa alpina*
3	7			*Rhinanthus minor*
	4			*Pyrola minor*
5	3		*Betula*	5
1	3		*Empetrum*	*sibiricum / subholarcticum*
1	3			*Trientalis europaea*
	1			*Crepis chrysantha*
4			*Rumex*	

ENG_A_3.2

Ligation	Fusion	Family	Genus	Species
79	841			*Bistorta vivipara*
89	457			*Viola biflora*
328	129			*Avenella flexuosa*
129	121			*Trientalis europaea*
15	42		*Viola*	*eppipsila / palustris*
128	33			*Phleum alpinum*
69	31		*Calamagrostis*	7
33	28		*Rumex*	5
8	18			*Rhinantus minor*
58	15	Poaceae	8	Poaceae 2
	2			*Cerastium fontanum*
2	2		*Deschampsia*	9
7	1		*Festuca*	5
1				*Poa alpina*
1			*Ranunculus*	

ENG_A_3.3

Ligation	Fusion	Family	Genus	Species
1564	869			*Avenella flexuosa*
407	120		*Festuca*	5
60	44		*Rumex*	5
3	17		*Ranunculus*	*acris / subborealis / turneri*
61	13			*Phleum alpinum*
	2		*Viola*	*eppipsila / palustris*
1	2			*Trientalis europaea*
part of 7	1	Poaceae	2	*Elymus* sp. / *Leymus* sp.
2	1	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
part of 7		Poaceae	8	Poaceae 2
1			*Deschampsia*	

ENG_C_1.1

Ligation	Fusion	Family	Genus	Species
284	1038	Rosaceae	2	*Rosa acicularis / Rubus sp.*
	75			*Bistorta vivipara*
24	73		*Rumex*	5
23	72	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
12	49		*Carex*	*brunnescens / gynocrates / parallela*
	33			*Trollius europaeus*
40	20			*Trientalis europaea*
2	1			*Comarum palustre*
	1			*Rubus idaeus*
1			*Pinus*	

ENG B 1.2

Ligation	Fusion	Family	Genus	Species
311	660		*Ranunculus*	*acris / subborealis / turneri*
264	395			*Avenella flexuosa*
261	281			*Anthoxanthum nipponicum*
14	198			*Viola biflora*
110	81			*Alchemilla glomerulans*
part of 19	46	Asteraceae	4	Asteraceae 2
12	32			*Bistorta vivipara*
2	27		*Equisetum*	*arvense / fluviatile / sylvaticum*
8	18			*Poa alpina*
11	11	Salicaceae		*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
28	9		*Calamagrostis*	7
12	9		*Deschampsia*	5
part of 19	4	Asteraceae	2	Asteraceae 6 : *Antennaria sp. / Omalotheca supina*
part of 19	3	Asteraceae	3	Asteraceae 5
11	2		*Empetrum*	*sibiricum / subholarcticum*
	2	Apiaceae	4	Liste Apiaceae
	1			*Omalotheca norvegica*
	1			*Pyrola minor*
1			*Luzula*	

ENG B 1.3

Ligation	Fusion	Family	Genus	Species
6	1			*Alchemilla glomerulans*
236	165			*Anthoxanthum nipponicum*
267	165			*Avenella flexuosa*
7	4			*Bistorta vivipara*
	1			*Crepis chrysantha*
24	10			*Poa alpina*
	1			*Pyrola minor*
26	162			*Viola biflora*
2			*Calamagrostis*	
82	38		*Deschampsia*	5
15	16		*Festuca*	9
2	4		*Luzula*	*multiflora / spicata / sudetica*
part of 918	474	Asteraceae	4	Asteraceae 2
part of 918	1	Asteraceae	3	Asteraceae 4
1		Poaceae		
	1	Rosaceae	2	*Rosa acicularis / Rubus sp.*
126	119	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*

ENG B 2.1

Ligation	Fusion	Family	Genus	Species
371	336			*Bistorta vivipara*
41	332			*Viola biflora*
168	140			*Avenella flexuosa*
99	106		*Ranunculus*	*acris / subborealis / turneri*
142	44			*Anthoxanthum nipponicum*
35	27		*Deschampsia*	5
25	20		*Equisetum*	*arvense / fluviatile / sylvaticum*
1	4			*Phleum alpinum*
2	4			*Poa alpina*
8	2	Poaceae	8	Poaceae 2
2	2	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
4	1			*Alchemilla glomerulans*
16			*Calamagrostis*	
6			*Festuca*	9

ENG C 1.2

Ligation	Fusion	Family	Genus	Species
667	488	Rosaceae	2	*Rosa acicularis / Rubus sp.*
183	313		*Rumex*	5
part of 377	250	Poaceae	8	Poaceae 2
	120			*Trientalis europaea*
56	96		*Carex*	*brunnescens / gynocrates / parallela*
	51			*Thymus oxyodontus*
part of 377	28	Poaceae	2	*Elymus* sp. / *Leymus* sp.
34	27		*Calamagrostis*	7
2	20			*Bistorta vivipara*
	19			*Avenella flexuosa*
14	11	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	8	Asteraceae	3	Asteraceae 4
	5	Asteraceae	4	Asteraceae 2
146	4			*Phlox hoodii*
	4		*Equisetum*	*palustre / pratense*
	3			*Anthoxanthum nipponicum*
	3		*Empetrum*	*sibiricum / subholarcticum*
	2		*Picea*	*glauca / mariana*
	2	Nymphaeaceae	2	*Nuphar pumila / Nymphaea tetragona*
1	1			*Beckmannia syzigachne*
	1			*Poa alpina*
	1			*Viola biflora*
	1		*Equisetum*	*arvense / fluviatile / sylvaticum*
1		Myrsinaceae		

ENG C 1.3

Ligation	Fusion	Family	Genus	Species
423	446			*Trientalis europaea*
65	118		*Carex*	*brunnescens / gynocrates / parallela*
64	67	Rosaceae	2	*Rosa acicularis / Rubus sp.*
	4		*Rumex*	5
	2	Poaceae	8	Poaceae 2
	1			*Avenella flexuosa*
1			*Pinus*	
17		Salicaceae		

ENG C 2.1

Ligation	Fusion	Family	Genus	Species
284	594			*Trientalis europaea*
267	316	Asteraceae	3	Asteraceae 4
42	229		*Equisetum*	*arvense / fluviatile / sylvaticum*
Part of 201	179	Poaceae	8	Poaceae 2
23	49			*Bistorta vivipara*
	16		*Rumex*	5
58	15		*Calamagrostis*	7
	7		*Carex*	*brunnescens / gynocrates / parallela*
11	5			*Anthoxanthum nipponicum*
1	4			*Poa alpina*
1	2		*Carex*	24
Part of 201	2	Poaceae	2	*Elymus* sp. / *Leymus* sp.
4	1			*Beckmannia syzigachne*
	1			*Carex lachenalii*
1	1	Rosaceae	2	*Rosa acicularis / Rubus sp.*

ENG_B_2.2

Ligation	Fusion	Family	Genus	Species
649	348			*Alchemilla glomerulans*
383	188			*Anthoxanthum nipponicum*
21	148			*Viola biflora*
305	120			*Avenella flexuosa*
70	92			*Bistorta vivipara*
11	82		*Equisetum*	*palustre / pratense*
66	82		*Ranunculus*	*acris / subborealis / turneri*
part of 79	33	Poaceae	8	Poaceae 2
13	12			*Poa alpina*
26	12		*Deschampsia*	5
15	11		*Festuca*	*9*
9	9	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
26	8		*Agrostis*	4
10	4			*Phleum alpinum*
	2		*Pinus*	*pumila / sylvestris*
part of 79	2	Poaceae	2	*Elymus* sp. / *Leymus* sp.

ENG_B_2.3

Ligation	Fusion	Family	Genus	Species
535	558			*Bistorta vivipara*
36	235			*Viola biflora*
510	218			*Avenella flexuosa*
404	172	Asteraceae	4	Asteraceae 2
173	64			*Anthoxanthum nipponicum*
61	48		*Deschampsia*	5
69	18		*Calamagrostis*	7
15	11	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
2	6			*Poa alpina*
5	6		*Festuca*	9
10	3	Poaceae	8	Poaceae 2
	1			*Nardus stricta*
1				*Crepis chrysantha*
1			*Carex*	
1			*Luzula*	
5			*Rumex*	

ENG_B_3.1

Ligation	Fusion	Family	Genus	Species
199	185		*Ranunculus*	*acris / subborealis / turneri*
411	158	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
273	153			*Avenella flexuosa*
27	79			*Viola biflora*
174	78		*Calamagrostis*	7
179	63			*Phleum alpinum*
135	57			*Anthoxanthum nipponicum*
41	18		*Deschampsia*	5
15	16			*Bistorta vivipara*
21	8		*Festuca*	9
8	5		*Rumex*	5
11	4			*Poa alpina*
	2	Poaceae	8	Poaceae 2
	2	Poaceae	2	*Elymus* sp. / *Leymus* sp.
	1		*Luzula*	*multiflora / spicata / sudetica*

ENG_C_2.2

Ligation	Fusion	Family	Genus	Species
422	727		*Carex*	*brunnescens / gynocrates / parallela*
873	646		*Calamagrostis*	7
48	82		*Rumex*	5
	78		*Pinus*	*pumila / sylvestris*
81	47			*Anthoxanthum nipponicum*
1	28			*Poa alpina*
	21	Poaceae	2	*Elymus* sp. / *Leymus* sp.
	14	Nymphaeaceae	2	*Nuphar pumila / Nymphaea tetragona*
	13			*Oxygraphis glacialis*
	8	Rosaceae	2	*Rosa acicularis / Rubus sp.*
4	3		*Equisetum*	*arvense / fluviatile / sylvaticum*
	2		*Deschampsia*	5

ENG_C_2.3

Ligation	Fusion	Family	Genus	Species
216	1334		*Equisetum*	*arvense / fluviatile / sylvaticum*
Part of 881	259	Asteraceae	3	Asteraceae 4
185	186		*Rumex*	5
148	143			*Bistorta vivipara*
192	47			*Anthoxanthum nipponicum*
19	19		*Carex*	*brunnescens / gynocrates / parallela*
	10			*Poa alpina*
8	6			*Trientalis europaea*
16	6	Rosaceae	2	*Rosa acicularis / Rubus sp.*
Part of 881	4	Asteraceae	2	Asteraceae 7 : *Aster alpinus / Erigeron sp.*
15	2			*Carex lachenalii*
	1			*Viola biflora*
Part of 881	1	Asteraceae	4	Asteraceae 2

ENG_C_3.1

Ligation	Fusion	Family	Genus	Species
1179	347		*Calamagrostis*	7
51	48			*Bistorta vivipara*
59	27		*Rumex*	5
30	16		*Luzula*	*multiflora / spicata / sudetica*
65	2			*Anthoxanthum nipponicum*
29	2			*Carex lachenalii*
1	1			*Poa alpina*
	1			*Viola biflora*
8				*Avenella flexuosa*

ENG_C_3.2

Ligation	Fusion	Family	Genus	Species
508	510			*Bistorta vivipara*
Part of 927	305	Poaceae	8	Poaceae 2
367	161		*Calamagrostis*	7
115	68		*Rumex*	5
98	26			*Anthoxanthum nipponicum*
15	24			*Poa alpina*
4	8			*Beckmannia syzigachne*
	6	Lycopodiaceae	2	*Diphasiastrum alpinum / Lycopodium sp.*
11	2			*Carex lachenalii*
	1			*Avenella flexuosa*
	1		*Empetrum*	*sibiricum / subholarcticum*
1	1		*Festuca*	9
	1	Asteraceae	3	Asteraceae 4
Part of 927	1	Poaceae	2	*Elymus* sp. / *Leymus* sp.

ENG B 3.2

Ligation	Fusion	Family	Genus	Species
1282	341			*Avenella flexuosa*
344	74			*Anthoxanthum nipponicum*
169	35		*Deschampsia*	5
16	34			*Viola biflora*
184	30		*Calamagrostis*	7
25	21			*Bistorta vivipara*
91	10	Poaceae	8	Poaceae 2
17	5			*Poa alpina*
14	4			*Phleum alpinum*
	3		*Rumex*	5
3	2			*Alchemilla glomerulans*
15	2	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
9	1		*Festuca*	9
1				*Beckmannia syzigachne*

ENG B 3.3

Ligation	Fusion	Family	Genus	Species
527	372	Asteraceae	4	Asteraceae 2
386	215			*Avenella flexuosa*
49	183		*Equisetum*	*palustre / pratense*
429	165			*Anthoxanthum nipponicum*
112	153			*Bistorta vivipara*
295	127		*Deschampsia*	5
73	89		*Rumex*	5
49	66		*Ranunculus*	*acris / subborealis / turneri*
13	38			*Viola biflora*
9	7		*Festuca*	9
9	5			*Poa alpina*
28	4		*Calamagrostis*	7
3	3	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
5	2			*Alchemilla glomerulans*
	2		*Carex*	*brunnescens / gynocrates / parallela*
4	1		*Luzula*	*multiflora / spicata / sudetica*
part of 2	1	Poaceae	8	Poaceae 2
part of 2	1	Poaceae	2	*Elymus* sp. / *Leymus* sp.
1			*Saussurea*	

ENG D 2.1

Ligation	Fusion	Family	Genus	Species
80	411			*Trollius europaeus*
37	249	Asteraceae	3	Asteraceae 4
69	196			*Bistorta vivipara*
53	181		*Festuca*	9
	103			*Viola biflora*
10	23			*Avenella flexuosa*
2	22	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
1?	6	Poaceae	2	*Koeleria asiatica / Trisetum spicatum*
1	5			*Anthoxanthum nipponicum*
	5		*Carex*	*brunnescens / gynocrates / parallela*
1	4			*Sibbaldia procumbens*
	2			*Poa alpina*
	1		*Equisetum*	*arvense / fluviatile / sylvaticum*
1?	1	Poaceae	8	Poaceae 2

ENG C 3.3

Ligation	Fusion	Family	Genus	Species
1891	922		*Calamagrostis*	7
245	150	Asteraceae	3	Asteraceae 4
85	125			*Bistorta vivipara*
176	64			*Anthoxanthum nipponicum*
65	51		*Rumex*	5
63	29	Poaceae	8	Poaceae 2
19	20			*Poa alpina*
4	18		*Carex*	*brunnescens / gynocrates / parallela*
7	13			*Carex lachenalii*
5	7		*Equisetum*	*arvense / fluviatile / sylvaticum*
1				*Beckmannia syzigachne*

ENG D 1.1

Ligation	Fusion	Family	Genus	Species
Part of 598	217	Asteraceae	3	Asteraceae 4
30	197			*Viola biflora*
833	193			*Trollius europaeus*
85	72		*Ranunculus*	*acris / subborealis / turneri*
92	35		*Festuca*	9
29	24	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
13	12			*Bistorta vivipara*
11	10			*Alchemilla glomerulans*
Part of 598	5	Asteraceae	4	Asteraceae 2
12	1			*Anthoxanthum nipponicum*
	1			*Pyrola minor*
	1	Apiaceae	4	

ENG D 1.2

Ligation	Fusion	Family	Genus	Species
40	438			*Viola biflora*
154	199		*Ranunculus*	*acris / subborealis / turneri*
72	100			*Bistorta vivipara*
Part of 125	65	Asteraceae	3	Asteraceae 4
164	45	Poaceae	2	*Koeleria asiatica / Trisetum spicatum*
50	14		*Festuca*	9
33	8			*Trollius europaeus*
1	1		*Luzula*	*multiflora / spicata / sudetica*
Part of 125	1	Asteraceae	4	Asteraceae 2
	1	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
1				*Sibbaldia procumbens*

ENG D 1.3

Ligation	Fusion	Family	Genus	Species
522	691			*Trollius europaeus*
96	260		*Ranunculus*	*acris / subborealis / turneri*
Part of 234	236	Asteraceae	4	Asteraceae 2
56	92			*Sibbaldia procumbens*
7	74			*Viola biflora*
28	42	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
6	26			*Bistorta vivipara*
	2			*Crepis chrysantha*
5	2		*Agrostis*	4
Part of 234	2	Asteraceae	3	Asteraceae 4
	1			*Avenella flexuosa*
	1		*Festuca*	9
11		Rosaceae		

HEI A 1.1

Ligation	Fusion	Family	Genus	Species
48	79	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
99	70			*Alchemilla glomerulans*
Part of 32	51	Asteraceae	4	Asteraceae 2
57	50			*Vaccinium uliginosum*
21	48			*Nardus stricta*
	45		*Ranunculus*	*acris / subborealis / turneri*
15	30		*Deschampsia*	5
5	22			*Bistorta vivipara*
	21		*Betula*	5
8	19			*Bartsia alpina*
Part of 32	18	Asteraceae	3	Asteraceae 4
	16			*Phleum alpinum*
	15			*Vaccinium alaskense*
43	11			*Cirsium heterophyllum*
	10		*Festuca*	9
	7			*Anthoxanthum nipponicum*
	7		*Agrostis*	4
	7		*Rumex*	5
	5			*Monolepis asiatica*
3	4		*Vaccinium*	*alaskense / myrtillus*
	3			*Trientalis europaea*
	1			*Chenopodium capitatum*
	1			*Geranium sylvaticum*
	1			*Urtica dioica*
	1			*Vaccinium myrtillus*
3			*Empetrum*	
1			*Epilobium*	
6			*Equisetum*	

HEI A 1.2

Ligation	Fusion	Family	Genus	Species
573	361			*Vaccinium uliginosum*
159	354		*Ranunculus*	*acris / subborealis / turneri*
608	223			*Alchemilla glomerulans*
178	160			*Nardus stricta*
77	56		*Deschampsia*	5
37	32		*Calamagrostis*	7
45	29		*Festuca*	9
31	23			*Bartsia alpina*
58	22	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
1	18			*Vaccinium alaskense*
2	17		*Agrostis*	4
	14			*Bistorta vivipara*
13	9			*Cirsium heterophyllum*
	5			*Geranium erianthum*
	4			*Euphrasia wettsteinii*
1	3			*Vaccinium myrtillus*
1	3		*Equisetum*	*arvense / fluviatile / sylvaticum*
	2			*Trientalis europaea*
5	2		*Betula*	5
24	2		*Vaccinium*	*alaskense / myrtillus*
	1			*Cornus suecica*
17				*Anthoxanthum nipponicum*
2			*Empetrum*	

ENG D 2.2

Ligation	Fusion	Family	Genus	Species
318	755			*Trollius europaeus*
150	212		*Festuca*	9
119	212	Asteraceae	3	Asteraceae 4
18	140		*Ranunculus*	*acris / subborealis / turneri*
56	127	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
72	126			*Avenella flexuosa*
12	38			*Bistorta vivipara*
3	34			*Viola biflora*
4	34		*Luzula*	*multiflora / spicata / sudetica*
Part of 2	15	Poaceae	2	*Koeleria asiatica / Trisetum spicatum*
4	5			*Alchemilla glomerulans*
Part of 2	2	Poaceae	8	Poaceae 2
	1		*Festuca*	*kolymensis / viviparoides*

ENG D 2.3

Ligation	Fusion	Family	Genus	Species
1	537			*Viola biflora*
14	277			*Trollius europaeus*
54	245		*Ranunculus*	*acris / subborealis / turneri*
Part of 37	103	Asteraceae	3	Asteraceae 4
8	57		*Festuca*	9
5	54			*Alchemilla glomerulans*
3	10			*Bistorta vivipara*
	3		*Carex*	*brunnescens / gynocrates / parallela*
	2			*Anthoxanthum nipponicum*
	2			*Avenella flexuosa*
	2		*Empetrum*	*sibiricum / subholarcticum*
	2		*Luzula*	*multiflora / spicata / sudetica*
Part of 37	2	Asteraceae	4	Asteraceae 2
	2	Poaceae	2	*Koeleria asiatica / Trisetum spicatum*
1	2	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	1			*Sibbaldia procumbens*

ENG D 3.1

Ligation	Fusion	Family	Genus	Species
89	725			*Viola biflora*
46	24			*Avenella flexuosa*
22	24		*Ranunculus*	*acris / subborealis / turneri*
6	14			*Anthoxanthum nipponicum*
24	14	Asteraceae	3	Asteraceae 4
15	12			*Bistorta vivipara*
39	11			*Trollius europaeus*
	3	Lycopodiaceae	2	*Diphasiastrum alpinum / Lycopodium sp.*
2	3	Rosaceae	2	*Sibbaldia procumbens / Sibbaldiopsis tridentata*
	1			*Geranium erianthum*
2	1			*Sibbaldia procumbens*
	1		*Cerastium*	6
	1		*Empetrum*	*sibiricum / subholarcticum*
1				*Poa alpina*
10			*Festuca*	

HEI_A_1.3

Ligation	Fusion	Family	Genus	Species
Part of 582	743	Asteraceae	3	Asteraceae 4
152	227			*Cirsium heterophyllum*
170	211			*Anthoxanthum nipponicum*
199	211			*Phleum alpinum*
121	156		*Festuca*	9
149	135			*Nardus stricta*
81	113		*Agrostis*	4
Part of 582	77	Asteraceae	4	Asteraceae 2
59	73		*Deschampsia*	5
47	71			*Vaccinium uliginosum*
18	60			*Bistorta vivipara*
10	49		*Ranunculus*	*acris / subborealis / turneri*
26	48		*Betula*	5
24	30			*Poa alpina*
4	25			*Geranium erianthum*
5	20			*Euphrasia wettsteinii*
12	6			*Alchemilla glomerulans*
16	5			*Cornus suecica*
1	4	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
1	3			*Bartsia alpina*
2	3		*Epilobium*	5
	2	Poaceae	8	Poaceae 2
2	1			*Geranium sylvaticum*
3	1			*Trollius europaeus*
2	1		*Vaccinium*	*alaskense / myrtillus*
2			*Calamagrostis*	

HEI_A_2.1

Ligation	Fusion	Family	Genus	Species
1311	924			*Avenella flexuosa*
	27			*Vaccinium alaskense*
	6			*Vaccinium myrtillus*
13	6		*Vaccinium*	*alaskense / myrtillus*
21	4		*Empetrum*	*sibiricum / subholarcticum*
2	3			*Alchemilla glomerulans*
	2	Poaceae	2	*Elymus* sp. / *Leymus* sp.
2	1			*Cornus suecica*
	1			*Vaccinium vitisidaea*
8			*Festuca*	

HEI_A_2.2

Ligation	Fusion	Family	Genus	Species
	123			*Vaccinium alaskense*
106	113			*Avenella flexuosa*
128	99		*Empetrum*	*sibiricum / subholarcticum*
29	12		*Vaccinium*	*alaskense / myrtillus*
1	8			*Vaccinium myrtillus*
3	2		*Betula*	5
	1			*Bistorta vivipara*
14	1			*Cornus suecica*
	1			*Vaccinium vitisidaea*

ENG_D_3.2

Ligation	Fusion	Family	Genus	Species
	494			*Trollius europaeus*
	358	Asteraceae	3	Asteraceae 4
	143		*Ranunculus*	*acris / subborealis / turneri*
	100			*Viola biflora*
	84			*Avenella flexuosa*
	15			*Bistorta vivipara*
	11		*Festuca*	9
	1	Poaceae	8	Poaceae 2

ENG_D_3.3

Ligation	Fusion	Family	Genus	Species
349	673			*Trollius europaeus*
107	110		*Festuca*	9
Part of 68	59	Asteraceae	3	Asteraceae 4
49	47			*Poa alpina*
20	46			*Bistorta vivipara*
44	39			*Alchemilla glomerulans*
1	19			*Viola biflora*
7	17			*Anthoxanthum nipponicum*
Part of 10	7	Poaceae	2	*Koeleria asiatica / Trisetum spicatum*
	3		*Carex*	24
Part of 10	3	Poaceae	8	Poaceae 2
	2		*Luzula*	*multiflora / spicata / sudetica*
	2	Lycopodiaceae	2	*Diphasiastrum alpinum / Lycopodium sp.*
	2	Rosaceae	2	*Sibbaldia procumbens / Sibbaldiopsis tridentata*
2	1			*Sibbaldia procumbens*
Part of 68	1	Asteraceae	2	Asteraceae 7 : *Aster alpinus / Erigeron sp.*

HEI_B_1.1

Ligation	Fusion	Family	Genus	Species
157	275		*Empetrum*	*sibiricum / subholarcticum*
32	79		*Betula*	5
85	71			*Avenella flexuosa*
14	38			*Bistorta vivipara*
27	32			*Vaccinium uliginosum*
	27		*Equisetum*	*arvense / fluviatile / sylvaticum*
	21			*Vaccinium alaskense*
51	18			*Phyllodoce caerulea*
46	13			*Cornus suecica*
34	11	Asteraceae	3	Asteraceae 4
	5	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	4			*Trollius europaeus*
	4			*Vaccinium myrtillus*
1	2			*Vaccinium vitisidaea*
	2		*Rhododendron*	*groenlandicum / tomentosum*
4			*Festuca*	
15			*Vaccinium*	

HEI_A_2.3

Ligation	Fusion	Family	Genus	Species
	122			*Vaccinium alaskense*
35	108			*Avenella flexuosa*
20	16		*Empetrum*	*sibiricum / subholarcticum*
1	14			*Vaccinium myrtillus*
36	13		*Vaccinium*	*alaskense / myrtillus*
	10			*Bistorta vivipara*
	10		*Potentilla*	15

HEI A 3.1

Ligation	Fusion	Family	Genus	Species
331	368			*Avenella flexuosa*
121	167		*Empetrum*	*sibiricum / subholarcticum*
	31			*Vaccinium alaskense*
2	10			*Vaccinium myrtillus*
	10	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	8	Poaceae	8	Poaceae 2
20	7		*Vaccinium*	*alaskense / myrtillus*
	6		*Pinus*	*pumila / sylvestris*
	5	Apiaceae	4	
14			*Festuca*	

HEI A 3.2

Ligation	Fusion	Family	Genus	Species
197	398		*Empetrum*	*sibiricum / subholarcticum*
8	67			*Avenella flexuosa*
	57	Apiaceae	4	
	49			*Vaccinium alaskense*
	11			*Vaccinium myrtillus*
7	5		*Vaccinium*	*alaskense / myrtillus*
6	1			*Cornus suecica*
	1			*Phyllodoce caerulea*
	1		*Ranunculus*	*acris / subborealis / turneri*

HEI A 3.3 (643 sequences)

Ligation	Fusion	Family	Genus	Species
199	247		*Empetrum*	*sibiricum / subholarcticum*
14	65			*Avenella flexuosa*
	21			*Vaccinium alaskense*
	19	Lycopodiaceae	2	*Diphasiastrum alpinum / Lycopodium sp.*
	15		*Festuca*	9
	6	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
9	3			*Cornus suecica*
	2			*Monolepis asiatica*
15	2		*Vaccinium*	*alaskense / myrtillus*
	1			*Vaccinium myrtillus*
1				*Vaccinium vitisidaea*

HEI C 1.1

Ligation	Fusion	Family	Genus	Species
	82			*Vaccinium alaskense*
7	45		*Betula*	5
	31			*Bistorta vivipara*
	13			*Alchemilla glomerulans*
79	12		*Empetrum*	*sibiricum / subholarcticum*
13	12		*Vaccinium*	*alaskense / myrtillus*
	12		*Rumex*	5
	9			*Avenella flexuosa*
	3			*Vaccinium vitisidaea*
	3	Rosaceae	2	*Rosa acicularis / Rubus sp.*
	1			*Vaccinium myrtillus*
31			*Pinus*	
9			*Puccinellia*	
1				*Stellaria borealis*

HEI B 1.2

Ligation	Fusion	Family	Genus	Species
353	359			*Vaccinium uliginosum*
61	102			*Trientalis europaea*
44	75			*Bistorta vivipara*
93	69		*Empetrum*	*sibiricum / subholarcticum*
239	27			*Cornus suecica*
	17			*Vaccinium alaskense*
9	16		*Betula*	5
22	9			*Avenella flexuosa*
16	6			*Anthoxanthum nipponicum*
	5			*Vaccinium myrtillus*
	5	Poaceae	2	*Elymus* sp. / *Leymus* sp.
	3		*Luzula*	*multiflora / spicata / sudetica*
	1		*Festuca*	9
	1	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
2				*Poa alpina*
9			*Vaccinium*	
2		Asteraceae		

HEI B 1.3

Ligation	Fusion	Family	Genus	Species
567	339			*Vaccinium uliginosum*
174	157		*Empetrum*	*sibiricum / subholarcticum*
91	80		*Betula*	5
129	19			*Cornus suecica*
24	13			*Avenella flexuosa*
	12			*Vaccinium alaskense*
6	5		*Luzula*	*multiflora / spicata / sudetica*
	4			*Bistorta vivipara*
	2			*Trollius europaeus*
1	1			*Trientalis europaea*
	1	Poaceae	2	*Elymus* sp. / *Leymus* sp.
	1		*Ranunculus*	*acris / subborealis / turneri*
15			*Vaccinium*	

HEI B 2.1

Ligation	Fusion	Family	Genus	Species
78	72			*Avenella flexuosa*
28	63		*Empetrum*	*sibiricum / subholarcticum*
	27	Apiaceae	4	
41	21			*Anthoxanthum nipponicum*
14	19			*Vaccinium uliginosum*
	9		*Betula*	5
	4			*Bistorta vivpara*
5	3			*Cornus suecica*
	1			*Phyllodoce caerulea*
	1			*Vaccinium alaskense*
	1			*Vaccinium vitisidaea*
	1		*Vaccinium*	*alaskense / myrtillus*
19		Salicaceae		

HEI_C_1.2

Ligation	Fusion	Family	Genus	Species
209	330		*Empetrum*	*sibiricum / subholarcticum*
41	108			*Avenella flexuosa*
1	62			*Vaccinium alaskense*
43	44		*Betula*	5
	5			*Vaccinium myrtillus*
	4			*Trientalis europaea*
36	3		*Vaccinium*	*alaskense / myrtillus*
	1			*Vaccinium vitisidaea*

HEI_C_1.3

Ligation	Fusion	Family	Genus	Species
528	357		*Empetrum*	*sibiricum / subholarcticum*
1	54			*Vaccinium alaskense*
44	7		*Betula*	5
	6			*Avenella flexuosa*
60	6		*Vaccinium*	*alaskense / myrtillus*
	2			*Vaccinium myrtillus*
	1			*Vaccinium vitisidaea*

HEI_C_2.1

Ligation	Fusion	Family	Genus	Species
353	160		*Empetrum*	*sibiricum / subholarcticum*
1	56			*Vaccinium alaskense*
71	9		*Vaccinium*	*alaskense / myrtillus*
1	5			*Vaccinium myrtillus*
	1			*Trollius europaeus*

HEI_C_2.2

Ligation	Fusion	Family	Genus	Species
	186		*Empetrum*	*sibiricum / subholarcticum*
	27			*Vaccinium alaskense*
	3		*Vaccinium*	*alaskense / myrtillus*
	2			*Vaccinium myrtillus*
	1	Asteraceae	4	Asteraceae 2
	1			*Trollius europaeus*

HEI_C_2.3

Ligation	Fusion	Family	Genus	Species
389	73		*Empetrum*	*sibiricum / subholarcticum*
1	65			*Vaccinium alaskense*
4	49			*Avenella flexuosa*
	34	Apiaceae	4	
	34	Poaceae	2	*Elymus* sp. / *Leymus* sp.
48	15		*Vaccinium*	*alaskense / myrtillus*
	10		*Deschampsia*	5
	8			*Vaccinium myrtillus*
	5			*Vaccinium vitisidaea*
	3	Asteraceae	3	Asteraceae 4
	3		*Equisetum*	*arvense / fluviatile / sylvaticum*
	2			*Bistorta vivipara*
	1		*Populus*	*tremula / tremuloides*
	1		*Rumex*	5

HEI_B_2.2

Ligation	Fusion	Family	Genus	Species
178	118			*Nardus stricta*
	11			*Vaccinium alaskense*
2	5			*Vaccinium myrtillus*
	2			*Cornus suecica*
	1	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
2	1		*Empetrum*	*sibiricum / subholarcticum*
1				*Phyllodoce caerulea*
12			*Betula*	
10			*Vaccinium*	

HEI_B_2.3

Ligation	Fusion	Family	Genus	Species
31	33		*Empetrum*	*sibiricum / subholarcticum*
41	24			*Avenella flexuosa*
1	18			*Vaccinium alaskense*
	6	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	4			*Vaccinium myrtillus*
	3			*Vaccinium vitisidaea*
	3		*Ranunculus*	*acris / subborealis / turneri*
	2			*Phyllodoce caerulea*
5			*Vaccinium*	

HEI_B_3.1

Ligation	Fusion	Family	Genus	Species
412	249		*Empetrum*	*sibiricum / subholarcticum*
29	45	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
22	24			*Vaccinium uliginosum*
	16			*Vaccinium alaskense*
99	15			*Cornus suecica*
10	15		*Betula*	5
16	14			*Avenella flexuosa*
	3			*Vaccinium myrtillus*
2	3		*Festuca*	9
11	3		*Vaccinium*	*alaskense / myrtillus*
	1		*Equisetum*	*arvense / fluviatile / sylvaticum*

HEI_B_3.2

Ligation	Fusion	Family	Genus	Species
316	677		*Empetrum*	*sibiricum / subholarcticum*
206	158	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
69	124			*Avenella flexuosa*
22	55		*Betula*	5
	25			*Vaccinium alaskense*
20	16			*Trientalis europaea*
72	13			*Cornus suecica*
	3			*Alchemilla glomerulans*
	3		*Ranunculus*	*acris / subborealis / turneri*
1	2			*Vaccinium myrtillus*
17	2		*Vaccinium*	*alaskense / myrtillus*
12	1		*Festuca*	9
1			*Carex*	

HEI C 3.1

Ligation	Fusion	Family	Genus	Species
365	473		*Empetrum*	*sibiricum / subholarcticum*
1	19		*Pinus*	*pumila / sylvestris*
	14			*Vaccinium alaskense*
21	4		*Vaccinium*	*alaskense / myrtillus*
	3			*Avenella flexuosa*
	2			*Vaccinium myrtillus*

HEI C 3.2

Ligation	Fusion	Family	Genus	Species
307	404		*Ranunculus*	*acris / subborealis / turneri*
53	251			*Viola biflora*
438	211			*Avenella flexuosa*
161	80		*Deschampsia*	5
52	31			*Phleum alpinum*
	11			*Bistorta vivipara*
29	6			*Anthoxanthum nipponicum*
8	2		*Festuca*	9
	1			*Trientalis europaea*
1	1		*Luzula*	*multiflora / spicata / sudetica*
6	1	Poaceae	8	Poaceae 2
10				*Botrychium lunaria*
1		Salicaceae		

HEI C 3.3

Ligation	Fusion	Family	Genus	Species
4	383			*Viola biflora*
255	328			*Avenella flexuosa*
41	61		*Deschampsia*	5
66	58		*Ranunculus*	*acris / subborealis / turneri*
56	36			*Anthoxanthum nipponicum*
19	33	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
29	26			*Bistorta vivipara*
3	21			*Trientalis europaea*
19	15		*Rumex*	5
8	9			*Phleum alpinum*
16	4		*Calamagrostis*	7
2	3			*Poa alpina*
1	2		*Festuca*	9
1?	2	Asteraceae	4	Asteraceae 2
1?	1	Asteraceae	3	Asteraceae 4
	1	Poaceae	8	Poaceae 2

HEI D 2.2

Ligation	Fusion	Family	Genus	Species
67	113		*Empetrum*	*sibiricum / subholarcticum*
	100			*Vaccinium alaskense*
19	51			*Trientalis europaea*
	20	Lycopodiaceae	2	*Diphasiastrum alpinum / Lycopodium sp.*
26	19	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	10			*Viola biflora*
14	10		*Vaccinium*	*alaskense / myrtillus*
	7			*Vaccinium myrtillus*
	1			*Anthoxanthum nipponicum*
	1			*Avenella flexuosa*
	1			*Bistorta vivipara*
16	1			*Cornus suecica*
	1		*Deschampsia*	5
	1		*Ranunculus*	*acris / subborealis / turneri*

HEI B 3.3

Ligation	Fusion	Family	Genus	Species
225	222		*Empetrum*	*sibiricum / subholarcticum*
47	49			*Vaccinium uliginosum*
83	9			*Cornus suecica*
4	8			*Bistorta vivipara*
8	7		*Betula*	5
3	4			*Avenella flexuosa*
	3			*Vaccinium alaskense*
	3		*Pinus*	*pumila / sylvestris*
1	1			*Phyllodoce caerulea*
	1	Rosaceae	2	*Rosa acicularis / Rubus sp.*
2				*Vaccinium vitisidaea*
3			*Vaccinium*	

HEI D 1.1

Ligation	Fusion	Family	Genus	Species
15	377			*Bistorta vivipara*
	162	Poaceae	2	*Elymus* sp. / *Leymus* sp.
	114	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	56		*Rumex*	5
	48	Apiaceae	4	
	44	Asteraceae	3	Asteraceae 4
	42		*Empetrum*	*sibiricum / subholarcticum*
	23			*Vaccinium alaskense*
	4			*Vaccinium myrtillus*
	2			*Phyllodoce caerulea*
6	2		*Vaccinium*	*alaskense / myrtillus*
	1			*Cornus suecica*
16			*Betula*	

HEI D 1.2

Ligation	Fusion	Family	Genus	Species
39	281			*Avenella flexuosa*
33	219	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	39	Lycopodiaceae	2	*Diphasiastrum alpinum / Lycopodium sp.*
	33			*Vaccinium alaskense*
	21		*Festuca*	9
31	19			*Cornus suecica*
2	4		*Vaccinium*	*alaskense / myrtillus*
1	3		*Betula*	5
1	2		*Carex*	24
	1			*Vaccinium myrtillus*
	1	Poaceae	8	Poaceae 2

HEI D 1.3

Ligation	Fusion	Family	Genus	Species
107	76		*Empetrum*	*sibiricum / subholarcticum*
33	63			*Avenella flexuosa*
1	47			*Vaccinium alaskense*
37	41	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	37	Lycopodiaceae	2	*Diphasiastrum alpinum / Lycopodium sp.*
9	23		*Betula*	5
41	9			*Cornus suecica*
14	9		*Vaccinium*	*alaskense / myrtillus*
	6			*Vaccinium myrtillus*
	6		*Pinus*	*pumila / sylvestris*
	4		*Ranunculus*	*acris / subborealis / turneri*
4	3		*Festuca*	9
	2			*Vaccinium vitisidaea*
	2	Asteraceae	3	Asteraceae 4
	1			*Bistorta vivipara*

HEI_D_2.1

Ligation	Fusion	Family	Genus	Species
315	143	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	84			*Vaccinium alaskense*
19	37			*Bistorta vivipara*
1	16			*Vaccinium myrtillus*
68	11		*Vaccinium*	*alaskense / myrtillus*
	10		*Huperzia*	*arctica / haleakalae*
	3		*Picea*	*glauca / mariana*
27	2			*Cornus suecica*
	1			*Trollius europaeus*
4	1		*Empetrum*	*sibiricum / subholarcticum*

HEI_D_2.3

Ligation	Fusion	Family	Genus	Species
179	226	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
120	180		*Empetrum*	*sibiricum / subholarcticum*
72	108		*Betula*	5
	53			*Vaccinium alaskense*
46	9			*Cornus suecica*
	6			*Vaccinium myrtillus*
18	2		*Vaccinium*	*alaskense / myrtillus*
6				*Avenella flexuosa*
1			*Huperzia*	
1		Lycopodiaceae		

HEI_D_3.1

Ligation	Fusion	Family	Genus	Species
	97			*Vaccinium alaskense*
	39	Poaceae	2	*Elymus* sp. / *Leymus* sp.
65	31		*Empetrum*	*sibiricum / subholarcticum*
	12			*Vaccinium myrtillus*
25	10		*Vaccinium*	*alaskense / myrtillus*
8				*Cornus suecica*
2			*Betula*	
4		Lycopodiaceae		

HEI_D_3.2

Ligation	Fusion	Family	Genus	Species
127	65		*Empetrum*	*sibiricum / subholarcticum*
	60			*Vaccinium alaskense*
13	13		*Vaccinium*	*alaskense / myrtillus*
3	6	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
	3			*Vaccinium myrtillus*
	1			*Vaccinium vitisidaea*
4				*Cornus suecica*

HEI_D_3.3

Ligation	Fusion	Family	Genus	Species
160	165		*Empetrum*	*sibiricum / subholarcticum*
1	89			*Vaccinium alaskense*
56	45	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*
15	27			*Avenella flexuosa*
61	16		*Vaccinium*	*alaskense / myrtillus*
	8			*Vaccinium myrtillus*
	6	Lycopodiaceae	2	*Diphasiastrum alpinum / Lycopodium sp.*
	2	Orchidaceae	3	

Appendix 3. Ambiguous Modern Plant Species

Agrostis
Agrostis exarata
Agrostis kudoi
Agrostis mertensii
Agrostis scabra
Apiaceae
Angelica archangelica
Angelica gmelinii
Conioselinum tataricum
Heracleum maximum
Podistera macounii
Asteraceae
Packera heterophylla
Packera hyperborealis
Packera ogotorukensis
Packera pauciflora
Senecio lugens
Senecio pseudoarnica
Asteraceae 2
Hieracium alpinum
Hieracium sect. Sylvatica
Lactuca sibirica
Saussurea alpina
Saussurea tilesii
Taraxacum croceum
Taraxacum lacerum
Asteraceae 3
Achillea alpina
Achillea millefolium
Arctanthemum arcticum
Artemisia arctica
Artemisia arctisibirica
Artemisia borealis
Artemisia frigida
Artemisia furcata
Artemisia glomerata
Artemisia kruhsiana
Artemisia tilesii
Chrysanthemum mongolicum
Tanacetum bipinnatum
Asteraceae 4
Eurybia sibirica
Solidago multiradiata
Solidago virgaurea
Symphyotrichum pygmaea
Asteraceae 5
Endocellion glaciale
Endocellion sibiricum
Petasites frigidus
Taraxacum alaskanum
Taraxacum arcticum
Taraxacum brachyceras
Taraxacum ceratophorum
Taraxacum holmenianum
Taraxacum lacerum
Taraxacum lateritium
Taraxacum phymatocarpum
Taraxacum soczavae

Asteraceae 6
Antennaria alpina
Antennaria friesiana
Antennaria monocephala
Antennaria rosea
Antennaria villifera
Omalotheca supina
Asteraceae 7
Aster alpinus
Erigeron acer
Erigeron borealis
Erigeron compositus
Erigeron eriocephalus
Erigeron humilis
Erigeron hyperboreus
Erigeron koraginensis
Erigeron porsildii
Erigeron purpuratus
Erigeron uniflorus
Betula
Betula fruticosa
Betula glandulosa
Betula nana
Betula neoalaskana
Betula pubescens
Calamagrostis
Calamagrostis canadensis
Calamagrostis deschampsioides
Calamagrostis inexpansa
Calamagrostis lapponica
Calamagrostis neglecta
Calamagrostis phragmitoides
Calamagrostis purpurascens
Carex
Carex aquatilis
Carex bigelowii
Carex capillaris
Carex fuliginosa
Carex glacialis
Carex holostoma
Carex krausei
Carex limosa
Carex lugens
Carex lyngbyei
Carex macrogyna
Carex media
Carex melanocarpa
Carex membranacea
Carex nigra
Carex norvegica
Carex petricosa
Carex podocarpa
Carex rariflora
Carex saxatilis
Carex scirpoidea
Carex stylosa
Carex subspathacea
Carex willamsii

Cerastium
Cerastium alpinum
Cerastium arcticum
Cerastium arvense
Cerastium beeringianum
Cerastium gorodkovianum
Cerastium regelii
Deschampsia
Deschampsia alpina
Deschampsia anadyrensis
Deschampsia brevifolia
Deschampsia cespitosa
Deschampsia sukatschewii
Dryas
Dryas ajanensis
Dryas drummondii
Dryas incisa
Dryas integrifolia
Dryas octopetala
Dryas punctata
Epilobium
Epilobium anagallidifolium
Epilobium davuricum
Epilobium hornemannii
Epilobium lactiflorum
Epilobium palustre
Eritrichium
Eritrichium aretioides
Eritrichium sericeum
Eritrichium splendens
Eritrichium villosum
Festuca
Festuca baffinensis
Festuca brachyphylla
Festuca brevissima
Festuca edlundiae
Festuca hyperborea
Festuca kolymensis
Festuca lenensis
Festuca ovina
Festuca rubra
Lycopodiaceae
Diphasiastrum alpinum
Lycopodium annotinum
Lycopodium clavatum
Oxytropis
Oxytropis arctica
Oxytropis borealis
Oxytropis bryophila
Oxytropis campestris
Oxytropis gorodkovii
Oxytropis jordalii
Oxytropis leucantha
Oxytropis maydelliana
Oxytropis mertensiana
Oxytropis middendorffii
Oxytropis sordida
Oxytropis varians

Orchidaceae
Amerorchis rotundifolia
Coeloglossum viride
Pseudorchis albida
Papaver
Papaver angustifolium
Papaver anjuicum
Papaver atrovirens
Papaver dahlianum
Papaver detritophilum
Papaver gorodkovii
Papaver lapponicum
Papaver macounii
Papaver mcconnellii
Papaver microcarpum
Papaver multiradiatum
Papaver nivale
Papaver nudicaule
Papaver paucistaminum
Papaver pulvinatum
Papaver radicatum
Papaver schamurinii
Papaver uschakovii
Papaver walpolei
Poaceae
Elymus alaskanus
Elymus fibrosus
Elymus macrourus
Elymus mutabilis
Elymus trachycaulus
Leymus arenarius
Leymus innovatus
Leymus interior
Leymus mollis
Poaceae 2
Alopecurus magellanicus
Anthoxanthum arcticum
Anthoxanthum monticola
Arctagrostis latifolia
Beckmannia syzigachne
Hyalopoa lanatiflora
Milium effusum
Poa abbreviata
Poa arctica
Poa eminens
Poa glauca
Poa hartzii
Poa laxa
Poa palustris
Poa paucispicula
Poa pratensis
Poa pseudoabbreviata
Poa sibirica
Vahlodea atropurpurea

Potentilla
Potentilla anachoretica
Potentilla arenosa
Potentilla crantzii
Potentilla crebridens
Potentilla elegans
Potentilla fragiformis
Potentilla hyparctica
Potentilla litoralis
Potentilla nivea
Potentilla pulchella
Potentilla rubella
Potentilla stipularis
Potentilla subvahliana
Potentilla uniflora
Potentilla villosula
Rosaceae
Rosa acicularis
Rubus acaulis
Rubus chamaemorus
Rubus stellatus
Rumex
Rumex acetosa
Rumex acetosella
Rumex arcticus
Rumex graminifolius
Rumex hultenii
***Salix* species**
Salix alaxensis
Salix arbusculoides
Salix arctica
Salix arctophila
Salix bebbiana
Salix chamissonis
Salix fuscescens
Salix glauca
Salix hastata
Salix herbacea
Salix jenisseensis
Salix lanata
Salix lapponum
Salix myrtilloides
Salix niphoclada
Salix nummularia
Salix ovalifolia
Salix phlebophylla
Salix phylicifolia
Salix polaris
Salix pulchra
Salix recurvigemmis
Salix reptans
Salix reticulata
Salix richardsonii
Salix rotundifolia
Salix saxatilis
Salix sphenophylla
Salix tschuktschorum
Salix uvaursi

Appendix 4. Plant Species identified from DNA extracts from Ice Bluff Cliff Permafrost Samples and their Bioclimatic Zones

EH288 (15810+-75)

Ligation	Fusion	Fus. 1a	Fus. 2a	Tot. Fus.	Family	Genus	Plant Species	Plant Bioclimatic Zone
	467	360	439	1266	Ranunculaceae		*Trollius europaeus*	Rare in D, frequent in E+N.
	241	251	250	742	Poaceae	*Festuca*	9	*V. capitata:*rare in C, frequent in D, E and N. *V. sambucifolia:* scattered in E, frequent in N. *V. transjenisensis:* rare in E, scattered in N.
250	248	204	225	677	Valerianaceae	*Valeriana*	*capitata / sambucifolia / transjenisensis*	
	199	197	226	622	Asteraceae	3	Asteraceae 4	
247	151	121	136	408	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*	
	81	76	68	225	Polygonaceae		*Bistorta vivipara*	Frequent in all zones (boreal to high arctic).
	58	84	81	223	Rosaceae		*Alchemilla glomerulans*	Scattered in zone C, frequent in D, E, and N.
	47	48	41	136	Pinaceae	*Larix*	*cajanderi / dahurica / laricina*	Scattered in E, frequent in N
	20	28	30	78	Ranunculaceae		*Anemone patens*	Rare in B, frequent in C-E+N.
		25	42	67	Poaceae		*Koeleria asiatica / Trisetum spicatum*	Rare in D, frequent in E+N.
	20	9	11	40	Asteraceae		*Hulteniella integrifolia*	
	6	6	18	30	Asteraceae		Asteraceae 3	Frequent in D, E and N
	4	5	4	13	Violaceae		*Viola biflora*	*V. myrtillus:* frequent in D, E and N
	1	2	2	5	Asteraceae	4	Asteraceae 2	Rare in D, scattered in E and N.
	1		3	4	Ericaceae		*Vaccinium alaskense*	
	1	1	1	3	Gentianaceae	*Gentianopsis*	*barbata / detonsa*	Boreal
		2		2	Asteraceae		*Omalotheca norvegica*	
	1			1	Ericaceae	*Vaccinium*	*alaskense / myrtillus*	
2					Cyperaceae	*Carex*		
2					Cyperaceae	*Kobresia*		

EH308* (20830+-90)

Ligation	Fusion				Family	Genus	Plant Species	Plant Bioclimatic Zone
part of 1273	314				Asteraceae		Asteraceae 3	*Achillea spp* and *A. arcticum:* frequent in D, E and N. *Artemisia spp.:* rare in C, scattered to frequent in other zones. *C. mongolicum* rare in D, scattered in E and frequent in N. *T. bipinnatum:* scattered in D, frequent in E and N.
	183				Asteraceae		*Hulteniella integrifolia*	Rare in B, frequent in C-E+N.
	37				Asteraceae		*Tanacetum bipinnatum*	Scattered in D, frequent in E and N.
	20				Boraginaceae	*Eritrichium*	4	Rare in A, scattered in B, rare to frequent in C-E and N.
	14				Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*	
	11				Rubiaceae	*Galium*	*ruthenicum / verum*	Both rare in E and scattered in N.
	7				Empetraceae	*Empetrum*	*sibiricum / subholarcticum*	
	5				Boraginaceae		*Myosotis alpestris*	Rare in zone A, frequent in all B-E+N.
part of 1273	4				Asteraceae		Asteraceae 6 : *Antennaria sp. / Omalotheca supina*	
	2				Papaveraceae	*Papaver*	19	
part of 1273	1				Asteraceae		Asteraceae 5	
	1				Betulaceae		*Alnus incana*	Rare in E, frequent in N
	1				Betulaceae		*Alnus viridis*	Scattered in E, frequent in N
	1				Poaceae		*Bromus pumpellianus*	Boreal-nemoral
181					Poaceae		9	
13					Plantaginaceae		*Plantago canescens*	Scattered in D and E, frequent in N.

EH316 (22960+-120)

Ligation	Fusion	Fus. 1a	Fus. 2a	Tot. Fus.	Family	Genus	Plant Species	Plant Bioclimatic Zone
	399	250	288	937	Polygonaceae		*Bistorta vivipara*	Frequent in all zones (boreal to high arctic).
	178	150	154	482	Equisetaceae	*Equisetum*	*arvense / fluviatile / sylvaticum*	*Equisetum spp.:* rare in zone A, rare to frequent in other zones.
	46	25	23	94	Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*	
	19	18	23	60	Plumbaginaceae		*Armeria scabra*	Rare in B, frequent in C-E, scattered in N.
	29	14	12	55	Lamiaceae		*Thymus oxyodontus*	Rare in zone D, frequent in E and N.
	14	12	17	43	Scrophulariaceae		*Lagotis glauca*	Scattered in A and N, frequent in B-E.
	14	13	10	37	Asteraceae	3	Asteraceae 4	
	18	8	9	35	Poaceae		*Avenella flexuosa*	Rare in zone D, frequent in E and N.
	10	8	9	27	Polygonaceae	*Aconogonon*	*alaskanum / ocreatum / tripterospermum*	*Aconogonum spp:* rare to frequent in D, frequent in E and N.
	11	5	10	26		*Rumex*	5	
	7	4	8	19	Asteraceae	2	*Packera* sp. / *Senecio* sp.	
	7	7	5	19	Poaceae		Poaceae 2	Hich arctic to boreal
	4	2	10	16	Ranunculaceae	*Ranunculus*	*acris / subborealis / turneri*	*R. acris:* rare in D, frequent in E and N. *R. subborealis:* rare in B, frequent in C-E+N. *R. turneri:* rare in D, E and N.
	9	3	3	15	Poaceae	*Festuca*	9	
	6	3	4	13	Asteraceae		*Hulteniella integrifolia*	Rare in B, frequent in C-E+N.
	4	3	4	11	Saxifragaceae		*Saxifraga hirculus*	Scattered in A, frequent in B-E+N
	7	1	1	9	Primulaceae		*Trientalis europaea*	Rare in D, frequent in E and N.
		4	4	8	Asteraceae	4	Asteraceae 2	
	3	3	1	7	Valerianaceae	*Valeriana*	*capitata / sambucifolia / transjenisensis*	*V. capitata:* rare in C, frequent in D, E and N. *V. sambucifolia:* scattered in E, frequent in N. *V. transjenisensis:* rare in E, scattered in N.
	5	1		6	Boraginaceae		*Myosotis alpestris*	Rare in zone A, frequent in all B-E+N.
	3	1	2	6	Asteraceae	5	Asteraceae 3	
	1	4	1	6	Empetraceae	*Empetrum*	*sibiricum / subholarcticum*	
	1	1	3	5	Poaceae		*Anthoxanthum nipponicum*	Rare in D, frequent in E and N.
		2	3	5	Asteraceae		*Crepis chrysantha*	Scattered in D, frequent in E and N.
		2	2	4	Saxifragaceae	*Saxifraga*	5	
	2		1	3	Papaveraceae	*Papaver*	19	
	2	1		3	Poaceae	2	*Elymus* sp. / *Leymus* sp.	*Elymus spp.* is rare in B and rare to frequent in the other zones. *Lymus* is rare in D and frequent in E + N.
	1	1	1	3	Ranunculaceae		*Trollius europaeus*	Rare in D, frequent in E+N.
	1	1	1	3	Poaceae	2	*Koeleria asiatica / Trisetum spicatum*	*K. asiatica* rare in D, frequent in E and N. *T.spicatum* rare in A, frequent in B-E+N.
	2			2	Scrophulariaceae		*Pedicularis oederi*	Rare in C, frequent in D, E and N.
	2			2	Violaceae		*Viola biflora*	Frequent in D, E and N
	1		1	2	Rosaceae		*Sanguisorba officinalis*	Rare in D, frequent in E and N.
	1	1		2	Ericaceae		*Vaccinium uliginosum*	Scattered in C, frequent in D, E and N.
	1		1	2	Scrophulariaceae	*Castilleja*	*elegans / hyperborea*	*C. elegans:* rare in C and N, frequent in D and E. *C. hyperborea:* rare in D and E, frequent in N.
	1		1	2	Rosaceae	*Potentilla*	15	
			2	2	Portulacaceae	*Claytonia*	*arctica / scammaniana*	*C. arctica:* rare in C and N, scattered in D and E. *C. scammaniana:* scattered in E and N.
			2	2	Poaceae	*Calamagrostis*	7	
	1			1	Ranunculaceae		*Anemone patens*	Rare in D, frequent in E+N.
	1			1	Poacae		*Beckmannia syzigachne*	Rare in E, frequent in N.
	1			1	Brassicaceae		*Cardamine pratensis*	Frequent in all zones (boreal to high arctic).
	1			1	Ericaceae		*Vaccinium myrtillus*	Frequent in D, E and N
		1		1	Poaceae		*Trisetum sibiricum*	Rare in D, frequent in E and N.
	1			1	Poaceae	*Deschampsia*	5	
	1			1	Ranunculaceae	*Thalictrum*	*alpinum / minus / sparsiflorum*	*Thalictrum spp.:* rare in C, rare to frequent in D, E and N
			1	1	Brassicaceae	*Parrya*	*arctica / nudicaulis*	*P. arctica:* scattered in A and E, frequent in B-D, rare in N. *P. nudicaulis:* rare in A, frequent in all other
			1	1	Ranunculaceae	*Caltha*	*arctica / palustris*	*C. palustris:* rare in D, frequent in E and N. *C. arctica:* rare in C, frequent in D, E and N.
			1	1	Fabaceae	*Astragalus*	*alpinus / umbellatus*	Both: rare in C, frequent in D, E and N.

EH287* (25440+-130)

Ligation	Fusion				Family	Genus	Plant Species	Plant Bioclimatic Zone
433	679				Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*	
56	63				Valerianaceae	*Valeriana*	*capitata / sambucifolia / transjenisensis*	*V. capitata:* rare in C, frequent in D, E and N. *V. sambucifolia:* scattered in E, frequent in N. *V. transjenisensis:* rare in E, scattered in N.
80	39				Rosaceae	*Dryas*	6	Dyas spp is rare in B and rare to frequent in the other zones.
9	20				Boraginaceae		*Myosotis alpestris*	Rare in zone A, frequent in all B-E+N.
	16				Saxifragaceae	*Saxifraga*	*nivalis / tenuis*	*S. nivalis:* scattered in the A, frequent B-E+N. *S. tenuis:* scattered in A and N, frequent in B-E.
19	13				Polygonaceae		*Bistorta vivipara*	Frequent in all zones (boreal to high arctic).
	7				Scrophulariaceae		*Lagotis glauca*	Scattered in A and N, frequent in B-E.
10	5				Ranunculaceae	*Caltha*	*arctica / palustris*	*C. palustris:* rare in D, frequent in E and N. *C. arctica:* rare in C, frequent in D, E and N.
	5				Apiaceae		*Angelica sylvestris / Cnidium cnidiifolium*	A. sylvestris: scattered in E, frequent in N. Cnidium cnidifolium rare in D, scattered in E, frequent in N.
	5				Papaveraceae	*Papaver*	19	
28	4				Poaceae	*Festuca*	9	
	4				Saxifragaceae	*Saxifraga*	*hieracifolia / nivalis / tenuis*	*S. hieracifolia:* rare in B, frequent in C-E, scattered in N.
	4				Fabaceae	*Oxytropis*	12	
4	3				Poaceae	2	*Elymus* sp. / *Leymus* sp.	*Elymus spp.* is rare in B and rare to frequent in the other zones. *Lymus spp.* is rare in D and frequent in E + N.
2	3				Saxifragaceae		*Saxifraga hirculus*	Scattered in A, frequent in B-E+N
	2				Ranunculaceae		*Anemone patens*	Rare in D, frequent in E+N.
1	2				Rosaceae	*Potentilla*	15	
5	1				Asteraceae	2	*Packera* sp. / *Senecio* sp.	
	1				Ranunculaceae		*Ranunculus arcticus*	Rare in B, frequent in C-E, scattered in N.
	1				Scrophulariaceae		*Pedicularis oederi*	Rare in C, frequent in D, E and N.
	1				Asteraceae		*Hulteniella integrifolia*	Rare in B, frequent in C-E+N.
2	1				Fabaceae		*Hedysarum hedysaroides*	Rare in B, frequent in C-E+N.
	1				Polygonaceae	*Rumex*	5	
1	1				Cyperaceae	*Carex*	24	
2					Caryophyllaceae		*Stellaria borealis*	Frequent in D,E and N.
1					Asteraceae		*Crepis chrysantha*	Scattered in D, frequent in E and N.
1					Poaceae		*Trisetum sibiricum*	Rare in D, scattered in E+N.
1					Betulaceae	*Betula*		Nemoral to zone E (rare in C).
1					Cyperaceae	*Kobresia*		

EH333 (28190+-160)

Ligation	Fusion				Family	Genus	Species	Plant Bioclimatic Zone
933	1072				Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*	
4	16				Cyperaceae	*Carex*	24	
23	9				Poaceae		Poaceae 2	
	4				Brassicaceae	*Parrya*	*arctica / nudicaulis*	*P. arctica:* scattered in A and E, frequent in B-D, rare in N. *P. nudicaulis:* rare in A, frequent in all other
13	3				Poaceae	*Calamagrostis*	7	*Calamagrostis spp.:* not in A, rare in B, scattered to frequent in other zones.
10					Poaceae	*Deschampsia*		
5					Valerianaceae	*Valeriana*		*Valeriana spp.:* rare in C, scattered to frequent in D, E and N.
2					Caryophyllaceae		*Stellaria borealis*	Frequent in D, E and N.
1					Pinaceae	*Picea*		Boreal

EH336 (29780+-210)

	Fusion				Family	Genus	Plant Species	Plant Bioclimatic Zone
	166				Equisetaceae	*Equisetum*	*arvense / fluviatile / sylvaticum*	*Equisetum spp.:* rare in zone A, rare to frequent in other zones.
	138				Polygonaceae		*Bistorta vivipara*	Frequent in all zones (boreal to high arctic).
	17				Poaceae		*Avenella flexuosa*	Rare in zone D, frequent in E and N.
	16				Salicaceae	3	*Salix sp. / Chosenia arbutifolia / Populus balsamifera*	
	9				Primulaceae		*Trientalis europaea*	Rare in D, frequent in E and N.
	9				Asteraceae		Asteraceae 4	
	8				Ranunculaceae		*Trollius europaeus*	Rare in D, frequent in E+N.
	7				Asteraceae		Asteraceae 2	
	5				Poaceae		*Anthoxanthum nipponicum*	Rare in D, frequent in E and N.
	3				Poaceae	*Deschampsia*	5	
	3				Empetraceae	*Empetrum*	*sibiricum / subholarcticum*	Check taxonomy with Brochmann
	3				Poaceae	*Festuca*	9	
	2				Rosaceae		*Alchemilla glomerulans*	Scattered in zone C, frequent in D, E, and N.
	2				Poaceae		*Poa alpina*	Rare in A, frequent in all other zones.
	2				Violaceae		*Viola biflora*	Frequent in D, E and N
	2				Ranunculaceae	*Ranunculus*	*acris / subborealis / turneri*	*R. acris:* rare in D, frequent in E and N. *R. subborealis:* rare in B, frequent in C-E+N. *R. turneri:* rare in D, E and N.
	2				Poaceae		Poaceae 2	
	1				Ranunculaceae		*Anemone patens*	Rare in D, frequent in E+N.
	1				Polygonaceae	*Rumex*	5	

*Mammoth DNA also recovered from this sample

Appendix 5. Ambiguous Permafrost Plant Species

Asteraceae
Packera heterophylla
Packera hyperborealis
Packera ogotorukensis
Packera pauciflora
Senecio lugens
Senecio pseudoarnica
Asteraceae 2
Hieracium alpinum
Hieracium sect. Sylvatica
Lactuca sibirica
Saussurea alpina
Saussurea tilesii
Taraxacum croceum
Taraxacum lacerum
Carex
Carex aquatilis
Carex bigelowii
Carex capillaris
Carex fuliginosa
Carex glacialis
Carex holostoma
Carex krausei
Carex limosa
Carex lugens
Carex lyngbyei
Carex macrogyna
Carex media
Carex melanocarpa
Carex membranacea
Carex nigra
Carex norvegica
Carex petricosa
Carex podocarpa
Carex rariflora
Carex saxatilis
Carex scirpoidea
Carex stylosa
Carex subspathacea
Carex willamsii
Poaceae
Elymus alaskanus
Elymus fibrosus
Elymus macrourus
Elymus mutabilis
Elymus trachycaulus
Leymus arenarius
Leymus innovatus
Leymus interior
Leymus mollis

Asteraceae 3
Achillea alpina
Achillea millefolium
Arctanthemum arcticum
Artemisia arctica
Artemisia arctisibirica
Artemisia borealis
Artemisia frigida
Artemisia furcata
Artemisia glomerata
Artemisia kruhsiana
Artemisia tilesii
Chrysanthemum mongolicum
Tanacetum bipinnatum
Calamagrostis
Calamagrostis canadensis
Calamagrostis deschampsioides
Calamagrostis inexpansa
Calamagrostis lapponica
Calamagrostis neglecta
Calamagrostis phragmitoides
Calamagrostis purpurascens
Deschampsia
Deschampsia alpina
Deschampsia cespitosa
Deschampsia brevifolia
Deschampsia anadyrensis
Deschampsia sukatschewii
Dryas
Dryas ajanensis
Dryas drummondii
Dryas incisa
Dryas integrifolia
Dryas octopetala
Dryas punctata
Eritrichium
Eritrichium aretioides
Eritrichium sericeum
Eritrichium splendens
Eritrichium villosum
Rumex
Rumex acetosa
Rumex acetosella
Rumex arcticus
Rumex graminifolius
Rumex hultenii

Asteraceae 4
Eurybia sibirica
Solidago multiradiata
Solidago virgaurea
Symphyotrichum pygmaea
Asteraceae 6
Antennaria alpina
Antennaria friesiana
Antennaria monocephala
Antennaria rosea
Antennaria villifera
Omalotheca supina
Festuca
Festuca baffinensis
Festuca brachyphylla
Festuca brevissima
Festuca edlundiae
Festuca hyperborea
Festuca kolymensis
Festuca lenensis
Festuca ovina
Festuca rubra
Oxytropis
Oxytropis arctica
Oxytropis borealis
Oxytropis bryophila
Oxytropis campestris
Oxytropis gorodkovii
Oxytropis jordalii
Oxytropis leucantha
Oxytropis maydelliana
Oxytropis mertensiana
Oxytropis middendorffii
Oxytropis sordida
Oxytropis varians
Potentilla
Potentilla anachoretica
Potentilla arenosa
Potentilla crantzii
Potentilla crebridens
Potentilla elegans
Potentilla fragiformis
Potentilla hyparctica
Potentilla litoralis
Potentilla nivea
Potentilla pulchella
Potentilla rubella
Potentilla stipularis
Potentilla subvahliana
Potentilla uniflora
Potentilla villosula

Asteraceae 5
Endocellion glaciale
Endocellion sibiricum
Petasites frigidus
Taraxacum alaskanum
Taraxacum arcticum
Taraxacum brachyceras
Taraxacum ceratophorum
Taraxacum holmenianum
Taraxacum lacerum
Taraxacum lateritium
Taraxacum phymatocarpum
Taraxacum soczavae
Papaver
Papaver angustifolium
Papaver anjuicum
Papaver atrovirens
Papaver dahlianum
Papaver detritophilum
Papaver gorodkovii
Papaver lapponicum
Papaver macounii
Papaver mcconnellii
Papaver microcarpum
Papaver multiradiatum
Papaver nivale
Papaver nudicaule
Papaver paucistaminum
Papaver pulvinatum
Papaver radicatum
Papaver schamurinii
Papaver uschakovii
Papaver walpolei
Poaceae 2
Alopecurus magellanicus
Anthoxanthum arcticum
Anthoxanthum monticola
Arctagrostis latifolia
Beckmannia syzigachne
Hyalopoa lanatiflora
Milium effusum
Poa abbreviata
Poa arctica
Poa eminens
Poa glauca
Poa hartzii
Poa laxa
Poa palustris
Poa paucispicula
Poa pratensis
Poa pseudoabbreviata
Poa sibirica
Vahlodea atropurpurea

Appendix 6

List of additional publications.

The effect of inappropriate calibration: Three case studies in molecular ecology. Ho, S., Saarma, U., Barnett, R., **Haile, J.,** Shapiro, B. (2008). PLoS ONE 3 (2), art. no. e1615

Ancient biomolecules from deep ice cores reveal a forested southern Greenland. *Willerslev, E., Cappellini, E., Boomsma, W., Nielsen, R., Hebsgaard, M.B., Brand, T.B., Hofreiter, M., Bunce, M., Poinar, H.N., Dahl-Jensen, D., Johnsen, S., Steffensen, J.P., Bennike, O., Schwenninger, J.L., Nathan, R., Armitage, S., de Hoog, C.J., Alfimov, V., Christl, M., Beer, J., Muscheler, R., Barker, J., Sharp, M., Penkman, K.E., **Haile, J.,** Taberlet, P., Gilbert, M.T., Casoli, A., Campani, E. and Collins, M.J. (2007). Science 317(5834): 111-4.*

Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. Binladen, J., Wiuf, C., Gilbert, M.T.P., Bunce, M., Barnett, R., Larson, G., Greenwood, A.D., **Haile, J.,** Ho, S.Y.W., Hansen, A.J. and Willerslev, E. (2006). Genetics 172(2): 733-741.

Reconstructing the world of the dodo. Rijsdijk, K.F., Hume, J.P., Bunnik, F., Vincent Florens, F.B., Baider, C., Shapiro, B., van der Plicht, H., Janoo, A., Griffiths, O., van den Hoek Ostende, L.W., Cremer, H., **Haile, J.,** Vernimmen, T. and Gittenberger, E. (in press). Quaternary Science Reviews.

Selection of a lectotype for Apteryx mantelli Bartlett, 1852 supported by genetic data. Cooper, J., **Haile, J.,** Scofield, P., Shepherd, L., Tennyson, A., Worthy, T. *In prep.*

Ancient DNA from insects. Thomsen, P., Elias, S., Gilbert, M., Munch, K., Sher, A., **Haile, J.,** Færch-Jensen, C., Willerslev, E. PloS ONE. *In press.*

The evolutionary history of the extinct ratite moa and New Zealand palaeogeography. Bunce, M., Worthy, T. H., Phillips, M., Holdaway, R.N., Willerslev, E., **Haile, J.,** Shapiro, B., Scofield, P., Drummond, A., Kamp, P., and Cooper A. *In prep.*