Canadian Journal of Zoology
Revue canadienne de zoologie

**Resurrecting phenotypes from ancient DNA sequences:
promises and perspectives**

Journal:	<i>Canadian Journal of Zoology</i>
Manuscript ID	cjz-2014-0337.R1
Manuscript Type:	Review
Date Submitted by the Author:	22-Jun-2015
Complete List of Authors:	Campbell, Kevin; University of Manitoba Hofreiter, Michael; University of Potsdam, Unit of General Zoology- Evolutionary Adaptive Genetics
Keyword:	paleophysiology, ancient DNA, extinct species, ADAPTATION < Discipline

Resurrecting phenotypes from ancient DNA sequences: promises and perspectives¹

K.L. Campbell² and M. Hofreiter³

¹This review is one of a series of invited papers arising from the symposium “Crossing boundaries and building bridges – integrative zoology”, sponsored by Canadian Science Publishing and held during the Genomes to Biomes joint conference, McGill University, 25–29 May 2014.

²Department of Biological Sciences, University of Manitoba, Winnipeg, R3T 2N2, Canada.
Email: Kevin.Campbell@umanitoba.ca

³Faculty of Mathematics and Life Sciences, Institute of Biochemistry and Biology, Unit of General Zoology–Evolutionary Adaptive Genomics, University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany. Email: michael.hofreiter@uni-potsdam.de

Corresponding author: Kevin L. Campbell, Department of Biological Sciences, 50 Sifton Road, University of Manitoba, Winnipeg, R3T 2N2, Canada. Tel: 204-474-6397. Fax: 204-474-7588. Email: Kevin.Campbell@umanitoba.ca

Resurrecting phenotypes from ancient DNA sequences: promises and perspectives

Kevin L. Campbell and Michael Hofreiter

Abstract: Anatomical changes in extinct mammalian lineages over evolutionary time, such as the loss of fingers and teeth and the rapid increase in body size that accompanied the late-Miocene dispersal of the progenitors of Steller's sea cows (*Hydrodamalis gigas* (Zimmermann, 1780)) into North Pacific waters, and the convergent development of a thick pelage and accompanying reductions in ear and tail surface area of woolly mammoths (*Mammuthus primigenius* (Blumenbach, 1799)) and woolly rhinoceros (*Coelodonta antiquitatis* (Blumenbach, 1799)), are prime examples of adaptive evolution underlying the exploitation of new habitats. It is likely, however, that biochemical specializations adopted during these evolutionary transitions were of similar or even greater biological importance. As these 'living' processes do not fossilize, direct information regarding the physiological attributes of extinct species has largely remained beyond the range of scientific inquiry. However, the ability to retrieve genomic sequences from ancient DNA samples, combined with ectopic expression systems, now permit the evolutionary origins and structural/functional properties of authentic prehistoric proteins to be examined in great detail. Exponential technical advances in ancient DNA retrieval, enrichment, and sequencing will soon permit targeted generation of complete genomes from hundreds of extinct species across the last one million years that, in combination with emerging *in vitro* expression, genome engineering, and cell differentiation techniques, promises to herald an exciting new trajectory of evolutionary research at the interface of biochemistry, genomics, palaeontology, and cell biology.

Key words: paleophysiology, ancient DNA, extinct species, adaptation, protein structure

48 **Our molecular past**

49 As famously lamented by Charles Darwin (1859), the paleontological record is depressingly
50 incomplete. This is, however, not unexpected as only a tiny fraction of carcasses ever become
51 fossils and only a tiny fraction of these are discovered by humans. Moreover, as only the hard
52 parts—in vertebrates mostly fragments of bones and teeth—tend to be preserved, much of
53 what we know about the biology, behaviour, and evolution of most extinct species is inferred
54 from partial remains. For a few species, for example woolly mammoths (*Mammuthus*
55 *primigenius*) and woolly rhinoceros (*Coelodonta antiquitatis*), much more detailed inferences
56 are gleaned from occasional finds—albeit much less common than we are led to believe—of
57 near complete mummified carcasses. Such discoveries have allowed stomach contents and
58 internal organs of extinct organisms to be examined together with visual phenotypes such as
59 the shape of ears or the length of tails and fur (e.g., Boeskorov et al. 2011; Fisher et al. 2012).
60 While the latter anatomical changes are prime examples of adaptive evolution underlying the
61 exploitation of Arctic habitats by these species, it is likely that physiological specializations
62 may have been of similar or greater importance during their adaptation to the harsh
63 environment of the Pleistocene Ice Ages. These dynamic processes are largely controlled by
64 proteins that, for example, form cellular channels, bind and transport ligands, or drive
65 chemical reactions; unfortunately, most are degraded quickly following death and become
66 completely lost during fossilization, precluding their analysis in long-dead species.

67 More than fifty years ago, however, Pauling and Zuckerkandl (1963) recognized that the
68 proteins of all living species contain a record of their historical past, and proposed that
69 ancestral sequence reconstruction could be used to deduce the primary structures of
70 polypeptides in the ancestors of contemporary species. Presciently, they moreover suggested
71 that via molecular “restoration studies” these archaic proteins could one day be synthesized
72 and used to “study the physico-chemical properties of these molecules and to make inferences
73 about their functions...[such as]...the oxygen affinity and its dependence on pH of ancestral
74 hemoglobins...as well as the affinity of ancestral enzymes for various substrates and the

probable nature of these substrates in past evolutionary history” (Pauling and Zuckerkandl 1963). This ambition was realized when two independent research groups used site-directed mutagenesis and protein engineering techniques to ‘resurrect’ and study the kinetic behaviour of a ribonuclease from a shared common ancestor of swamp buffalo, river buffalo, and ox that lived ~5 million years ago (Ma) (Stackhouse et al. 1990), and the activity, thermostability, and three-dimensional structure of lysozymes from a game bird progenitor in the order *Galliformes* (Malcolm et al. 1990). Similar phylogenetically informed biotechnological approaches have subsequently been used to pinpoint the timing and molecular underpinnings of key adaptive functional changes in ancient proteins, including pancreatic ribonucleases in leaf-eating monkeys (Zhang 2006), visual pigments (rhodopsins) from archosaurs and early vertebrates (Chang et al. 2002; Shi and Yokoyama 2003), a primordial hormone receptor from a Proterozoic protostome-deuterostome common ancestor (Thornton et al. 2003), and even an elongation factor protein from 0.5 to 3.5 billion year old forbearers of modern bacteria (Gaucher et al. 2008).

Although a strength of this methodology is that it allows for the testing of hypotheses regarding molecular adaptations in deep evolutionary time (e.g., Jermann et al. 1995; Thornton 2004), it is unable to examine lineages with no living descendants, such as the many recently extinct Ice Age species which presumably evolved under strong selective pressures (Campbell et al. 2010). Notably, these species represent a substantial fraction of recent mammalian diversity, as large-scale mass extinction events over the past 40,000 years annihilated ~80% of large-bodied (>100 kg) North American and ~35% of northern Eurasian mammalian species (Stuart 1991). Herein we review how combining technical advances in ectopic expression systems with methodologies that permit the retrieval of nuclear gene and protein sequences from ancient biological remains has not only opened an exciting new chapter in palaeobiological research, but promises to add a new dimension to the study of natural selection of both extinct and extant species.

Extinct species as exemplars of the August Krogh Principle

Given their small ears, short tail, stocky legs, thick woolly undercoat, and intimate association with frigid northern environments, it is a surprise for many to learn that the ancestors of woolly mammoths not only originated in sub-tropical Africa some 6.7 Ma, but that they are more closely related to Asian elephants (*Elaphas maximus* L., 1758) than either species are to the African elephants (*Loxodonta africana* (Blumenbach, 1797) and *L. cyclotis* (Matschie, 1900)) (Rohland et al. 2007, 2010). Moreover, it is clear from the fossil record that the first ancestral mammoths only entered northern Siberia ≤ 2 Ma (Lister et al. 2005), at the dawn of one of the most profound global cooling events in the Earth's history, the Pleistocene Ice Ages. Hence, in contrast to their almost certainly sparsely haired tropical ancestors (and their extant Asian and African relatives), the above noted morphological specializations indicate that a strong selection pressure faced by early Siberian mammoths regarded the conservation of metabolic heat. It thus stands to reason that woolly mammoths, together with other recently extinct species that independently evolved to exploit cold Arctic and sub-Arctic environments, such as woolly rhinoceros and Steller's sea cows (*Hydrodamalis gigas*), provide particularly attractive model systems to study physiological adaptations to cold climates. Indeed, each lineage exhibits a relatively low molecular evolution rate (Meredith et al. 2011), that, together with relatively recent divergence dates from their closest living ancestors (e.g., 17.5–22.8 Ma for the woolly rhino/Sumatran rhino split; Willerslev et al. 2009), should facilitate the identification of key amino acid replacements underlying functional adaptation as the accumulation of additional random changes are expected to be minimal. Finally, each group is of an easily manageable size for comparative phylogenetic analyses (i.e., simple three species comparisons; Garland and Adolph 1994) and for tests of convergent evolution. Taken together, these features align nicely with the August Krogh Principle, which states: “for such a large number of problems there will be some animal of choice or a few such animals on which it can be most conveniently studied” (Krogh 1929). Extant high-Arctic mammals (e.g., musk-ox, reindeer, and polar bears), by comparison, do not

have a clear evolutionary history of recent movement from tropical to Arctic environments, nor do they have non-temperate sister taxa for phylogenetically informative three-species comparisons. However, despite the advent of a bacterial expression system for the ectopic synthesis of proteins in the late 1970's (Itakura et al. 1977), functional studies on woolly rhinos, mammoths, and Steller's sea cows (together with a host of other recently extinct 'August Krogh Principle' exemplar species for addressing separate biological questions; Table 1) were previously hindered by an inability to obtain nucleotide sequence data from these species.

Ancient DNA

The sole purpose of numerous processes in the bodies of all living organisms is to constantly repair the organic molecules (e.g., DNA, proteins, carbohydrates, and lipids) from which they are made. However, immediately following death, endogenous nucleases together with an army of invading organisms (predominantly bacteria and fungi), begin the inexorable degeneration of organic molecules within the body, leading to the near-complete abolishment of organic matter within a few decades (Hofreiter et al. 2001). However, upon desiccation or following freezing (i.e., in permafrost), the rate of degradation may be dramatically slowed, pushing the envelope back many thousands of years. Regardless, even under the best circumstances, DNA extractions from ancient specimens tend to be heavily contaminated with foreign DNA from invading microorganisms and contain only relatively small amounts of endogenous DNA. Hydrolytic damage moreover degrades these remaining molecules into short fragments—often 100 base pairs (bp) or less (Poinar et al. 2006; Green et al. 2010; Dabney et al. 2013)—that are fraught with chemical modifications arising from radiation, oxidation, and deamination (Lindahl 1993; Hofreiter et al. 2001).

It is thus impressive that the field of ancient DNA was already established 30 years ago (prior to the advent of the polymerase chain reaction or PCR), via the successful extraction, cloning, and sequencing of two small fragments (totaling 229 bp) of endogenous DNA from

140 year old dried muscle samples from an extinct sub-species of zebra (quagga; *Equus quagga quagga* (Boddaert, 1785)) (Higuchi et al. 1984). However, this study and the many that followed over the next 15 years concentrated on the small circular mitochondrial DNA molecules that are generally present in hundreds to thousands of copies per cell, but accounts for only a minute fraction (~16,500 bp or <0.00005%) of all genetic material in mammals. It also encodes only a handful of mitochondrial proteins, precluding its applicability for functional studies. It wasn't until 1999 that the use of dedicated ancient DNA facilities and strict adherence to procedures that minimize extraneous DNA contamination from modern sources demonstrated that small fragments (43 to 58 bp) of Pleistocene nuclear DNA—present in only two copies per cell—can survive in amounts sufficient for PCR amplification in at least some permafrost remains (Greenwood et al. 1999). Although these nuclear gene sequences provided little to no information about the structure and function of the ancient proteins they encoded, this work represented an important early advance in the development of the field.

Linking genotype with phenotype

With rare exceptions, such as those involving pseudogenization of tooth-specific (e.g., enamel; Meredith et al. 2009; Springer et al. 2015) or eye-specific genes (e.g., retinal opsins; Meredith et al. 2013), our ability to ascribe a genotype to a specific phenotypic variant in extinct organisms is virtually non-existent (Shapiro and Hofreiter 2014). However, the ability to obtain ancient nuclear DNA sequences, together with the ability to create and ectopically express recombinant proteins, in theory opened the door to the genetic reconstruction of extinct phenotypes from sub-fossil material. It nonetheless remained largely impractical to sequence the hundreds to thousands of nucleotides—often spanning multiple coding exons—that comprise most genes until the development of two-step multiplex approach that allowed for the simultaneous amplification of the first complete Ice Age mitochondrial genome (not surprisingly, a woolly mammoth) in two reactions (Krause et al.

2006). This advance quickly lead to the retrieval of the first complete nuclear encoded gene, the melanocortin 1 receptor (*MC1R*; 1236 bp), from an extinct species: again, from the woolly mammoth (Römpler et al. 2006). This gene is one of a large number of genes known to be involved in colour determination of both birds and mammals (Rees 2003), and was examined to test whether differences in hair coloration—from blond to red to almost black—of permafrost-preserved mammoths could be explained by variations in gene function as opposed to chemical factors in the sediments to which the hairs had been exposed for tens of thousands of years (see e.g., Tridico et al. 2014). A follow-up study focused on a short (128 bp) fragment of this locus from two Neanderthal specimens (Lalueza-Fox et al. 2007). Notably, both studies revealed heterozygous alleles with one copy exhibiting unique amino acid exchanges not found in living elephants or humans, respectively. *In vitro* functional analyses of proteins harbouring these amino exchanges expressed in an immortalized line of monkey kidney cells further demonstrated sharp reductions in both basal activity and responses to two agonists, aMSH and NDP-aMSH, in the derived mammoth and Neanderthal alleles suggesting the presence of blond and fair-skinned/red-head phenotypes, respectively (Römpler et al. 2006; Lalueza-Fox et al. 2007). However, for mammoths at least, single nucleotide polymorphism typing of *MC1R* fragments from 47 additional mammoths revealed the presence of only a single heterozygous individual carrying the derived allele, indicating that light-coloured mammoths, if present in the population, were extremely rare (Workman et al. 2011). Nonetheless, these pioneering studies demonstrated that the genetic reconstruction of phenotypes from extinct species had become a practical reality.

Another highly conspicuous externally visible characteristic of woolly mammoths is their extremely thick overcoat that carried hairs up to a meter in length. As similar (though less dramatic) long-hair phenotypes in mice, dogs, and cats are associated with non-functionalization of the fibroblast growth factor 5 (*FGF5*) gene, this locus together with four keratin genes were investigated (Roca et al. 2009). However, only synonymous (silent) mutations were identified on the mammoth *FGF5* locus, while the mammoth keratin genes

did not possess any unique substitutions, indicating that other loci must underlie the thick woolly phenotype of mammoths (Roca et al. 2009). In this regard, 38 candidate genes that may underlie both coat colour and hair morphology in this lineage were identified in a recent study that sequenced the genomes of two woolly mammoths to ~20-fold coverage (Lynch et al. 2015). These authors also revealed mammoth-specific residue exchanges in four transient receptor potential (TRP) channels—TRPA1, TRPM4, TRPV3, and TRPV4—implicated in temperature discrimination in mammals (Lynch et al. 2015). Computer modelling of these exchanges onto cryo-electron microscopic structures moreover pinpointed residues on the TRPA1 and TRPV4 proteins that are predicted to alter their gating properties (i.e., maximal activity and range of temperatures over which they are activated). Intriguingly, a woolly mammoth mutation (N647D) located in a loop region of TRPV3 that forms an ion pore for Ca^{2+} was predicted to have similar effects and has been shown to abolish heat sensitivity in the orthologous mouse protein (Grandl et al. 2008). Briefly, this heat activated cation channel is expressed in skin, sensory neurons, and epidermal keratinocytes and is stimulated at temperatures $>33^{\circ}\text{C}$ in mice (Moqrich et al. 2005). Intriguingly, this channel also co-localizes with hair follicles, where it further operates as a regulator of hair morphogenesis (Cheng et al. 2010). Functional expression of the mammoth and inferred ancestral Asian elephant/mammoth TRPV3 channels in HEK293 cells revealed that while the activation temperatures ($\sim 29^{\circ}\text{C}$) of both proteins were similar, the maximal activity of mammoth TRPV3 was reduced by ~20% relative to that of the ancestral elephantid Ca^{2+} channel (Lynch et al. 2015). As TRPV3 activation inhibits hair shaft elongation and hair follicle regression in mice, the authors accordingly suggest the N647D substitution may in part underlie the evolution of a woolly phenotype in mammoths. Phenotypic characterization of genes relating to hair cover and/or thermal sensation in other cold-adapted species (e.g., woolly rhinos, Steller's sea cow) have yet to be examined and offer a fertile avenue for research.

As noted earlier, morphological adaptations often play key roles in a species' adaptation to new environments. However, in certain instances these modifications require the presence of

specific biological/physiological traits prior to or concurrent with their evolution. One example is the counter-current heat exchangers of regionally heterothermic vertebrates, which allow for large temperature gradients between the core and extremities. For cold adapted mammals, this vascular arrangement reduces heat loss to the environment thereby lowering daily energy requirements when thermal costs are at their highest, and forage quality and availability is often at its lowest (i.e., during winter). Unlike tropical species, such as humans, which possess blood with a high thermal sensitivity (whereby small increases in temperature cause relatively large increases in oxygen offloading)—ostensibly to promote increased oxygen delivery to warm exercising muscles—temperate species appear to require blood that exhibits a relatively low thermal sensitivity to maintain appropriate oxygen delivery to cold tissues (Weber and Campbell, 2011). This altered blood phenotype arises from amino acid substitutions in the hemoglobin molecules that inherently lower the energetic transition of the protein from the Tense (deoxygenated) to Relaxed (oxygenated) state (Weber et al. 2014), and/or via the binding of other molecules inside the red blood cells (primarily protons, chloride ions, and organic phosphates such as 2,3-diphosphoglycerate) to hemoglobin which release heat, which can be used to help break the heme–oxygen bond (Weber and Campbell 2011). In the latter case, it should be noted that hemoglobins of non-cold-adapted species may also bind these same effector molecules; ‘cold-adapted’ hemoglobins simply possess more effector binding sites and thus their oxygen affinity is less sensitive to changes in temperature.

Given the arctic habitat of woolly mammoths, it was hypothesized that they evolved substitutions in their hemoglobin that may, at least in part, underlie their physiological adaptation to the cold (Campbell et al. 2010). DNA sequencing revealed that the mammoth β -type (β/δ) globin chain only differed from Asian elephants at 3 of 146 amino acid positions, while the α -type chain had not acquired any substitutions following the split of these two lineages. As predicted from computer modelling, the oxygenation properties of the recombinant mammoth protein measured in the presence and absence of red blood cell effector molecules and at various physiologically relevant temperatures revealed radically

different functional attributes relative to those of Asian elephants (Campbell et al. 2010; Yuan et al. 2011). The most striking change was a strong increase in the inherent oxygen affinity of the molecule likely caused by a charge changing residue exchange ($\beta/\delta 101\text{Glu}\rightarrow\text{Gln}$) within the sliding interface of the protein. Remarkably, however, this same amino acid substitution also appeared to establish additional effector binding sites for chloride and protons that not only negated this deleterious effect, but also significantly lowered the proteins' sensitivity to temperature (Campbell et al. 2010; Yuan et al. 2011). From an evolutionary standpoint it bears mentioning that no other amino acid replacement at this residue position appears to be able to create this new binding site, indicating that mammoths evolved the only possible beneficial mutation at this position. These conclusions were corroborated by X-ray crystallography and solution NMR studies (Yuan et al. 2011; Noguchi et al. 2012), coupled with site-directed mutagenic experiments whereby individual and paired mammoth changes were introduced into the Asian elephant background (Yuan et al. 2013), that permitted the structural and mechanistic dissection of this extinct physiological phenotype. Such investigations are required to elucidate the functional role of novel amino acid replacements not found in living species or for pinpointing the relative contributions of individual mutations, and are likely to be employed more in future as more archaic proteins are resurrected.

***In vivo* expression**

While amino acid polymorphisms and gene inactivation presumably drive many phenotypic differences observed in extinct species relative to their extant relatives, variations in noncoding *cis*-regulatory elements that govern the location, timing, or level of expression of functional loci may also underlie important phenotypic innovations (Wray 2007). Unfortunately, relatively few of these regions have been identified and characterized in model taxa, let alone extinct species. Early progress in this latter area was achieved by generating an expression construct containing the transcriptional enhancer element of the Tasmanian tiger

(*Thylacine cynocephalus* (Harris, 1808)) *pro α 1(II)* collagen (*Col2a1*) gene with the human β -globin basal promoter flanked by a *lacZ* (β -galactosidase) reporter gene. Histological staining two-weeks after injection of this construct into mouse zygotes resulted in chondrocyte-specific expression in the developing fetuses. Although the nucleotide differences in the thylacine enhancer compared to the mouse version did not alter the function of this element (Pask et al. 2008), this work was the first to illustrate the future promise of elucidating variations in the function of regulatory sequences from extinct species in an *in vivo* model system.

More recently, an *in vivo* system was also used to investigate the functionality of the *tbx5* gene in extinct, New Zealand endemic moas. The transcription factor encoded by this locus is known to play a critical role in forelimb development in all vertebrates (Ng et al. 2002; Agarwal et al. 2003). Moas show the most extreme reduction of the forelimb skeleton among ratites, a (paraphyletic) group of flightless birds from the southern hemisphere, though, phylogenetically, they form the sister group to the fully-flighted South American tinamous (Phillips et al. 2010). Briefly, Huynen et al. (2014) investigated the moa *tbx5* sequence structure and transcription activity in both cell culture and chicken embryo assays. Surprisingly, and despite wingless moa likely having lost their entire forelimb girdles >20 Ma (save for a small residual shoulder bone, the scapulocoracoid), their *tbx5* gene not only showed transcription activation similar to that of the chicken homolog in cell culture assays, but it also triggered the development of forelimb features when injected into the hindlimb of chicken embryos (Huynen et al. 2014). These results demonstrate that while the moa *tbx5* gene presumably maintains a role in initiating forelimb/pectoral girdle (scapulocoracoid) development, it is not sufficient for continued forelimb growth (Huynen et al. 2014).

Another example explored the effects of two amino acid changes in a gene (*FOXP2*) critical for the development of speech and language in humans using a mouse model system. Earlier studies identified two residue replacements in the human protein that arose following the split of humans from chimpanzees (5–10 Ma), but prior to the divergence of humans from

other archaic (i.e., Neanderthal and Denisovan) human lineages (Enard et al. 2002, Krause et al. 2007). Introduction of these ‘humanized’ changes into mice altered neuronal properties of basal ganglia that affect ultrasonic vocalizations, suggesting an important role in the evolution of human vocal communication (Enard et al. 2009). More recently, a nucleotide substitution in a highly conserved region of intron 8 of *FOXP2* conserved in most (98%) modern humans, but not found in other anthropoid primates, including Neanderthals and Denisovans, was revealed to affect a binding site for another transcription factor (*POU3F2*) exclusively expressed in the central nervous system (Maricic et al. 2013). Intriguingly, *in vitro* functional assays further revealed that the ancestral (pre-modern human) version was more efficient in activating transcription from a reporter construct, suggesting a prominent role for this mutation in a recent selective sweep of this locus in modern humans (Maricic et al. 2013). However, studies employing transgenic mice that carried each of the *POU3F2* variants fused to a *lacZ* reporter failed to detect enhancer activity, suggesting that the archaic and modern 1.8-kb-long human sequences do not work in the mouse genetic background (Maricic et al. 2013). Despite this setback, similar *in vivo* approaches are likely to be central to deciphering the roles of derived mutations in regulatory elements or complex biochemical pathways of extinct species in the future.

Finally, novel insights into the physiological attributes of extinct species have also been achieved via inclusion of specific gene products into ancestral sequence reconstruction datasets of proteins known to strongly influence phenotypic traits. The oxygen-storing muscle pigment myoglobin, for example, plays a key role in the ability of air-breathing birds and mammals to hold their breath and actively forage under water. Using myoglobin sequence data from 128 extant and two extinct species (woolly mammoth and Steller’s sea cow), Mirceta et al. (2013) revealed a molecular signature of increased surface protein charge in all lineages of diving mammals with an extended aquatic history—from 15 g water shrews to 80,000 kg whales—that underlies this phenotype via its mechanistic linkage with muscle myoglobin content and hence maximal dive time. Apart from providing insights into the

tempo and multiple convergent routes of dive capacity evolution across the mammalian phylogeny, the ancestral sequence reconstruction approach employed by these authors allowed for a comprehensive modelling of the muscle oxygen storage capacity and maximal submergence times—figuratively putting ‘flesh onto the bones’—of successive transitional forms of whales, sea cows, and seals over more than 50 million years of evolution. Importantly, inclusion of woolly mammoths and Steller’s sea cow into this dataset substantially increased the sample size, and hence reliability of ancestral sequence reconstruction for the afrotherian clade that provided unique evidence for a shared early Palaeocene amphibious ancestor of sea cows, elephants, and hyraxes (Mirceta et al. 2013). The ability to trace protein evolution over geological time scales further permitted these authors to deduce that the longer maximal submergence time observed for Steller’s sea cows (~300 sec) by Georg Wilhelm Steller (Steller 1751) relative to extant sirenians primarily arose from massive increases in body size (up to 11,000 kg) in the former lineage as opposed to physiological specializations. As sequence data from Holocene/Pleistocene organisms becomes more readily available we anticipate that similar approaches will allow the evolution and characterization of other phenotypic traits to be reconstructed for extinct species in unprecedented detail.

Future perspectives

To date, functional ancient DNA studies have focused on one or only a couple of loci, yet many morphological and physiological phenotypes are determined by a complex interaction of genes. Unfortunately, even the most efficient multiplex PCR techniques are quite laborious and thus PCR is not really an option to study the myriad of genes potentially involved in, for example, the long hair phenotype of woolly mammoths and woolly rhinoceros. However, the development of specifically tailored ancient DNA extraction methodologies that optimize recovery of the miniscule amounts of DNA within prehistoric samples (Dabney et al. 2013), together with parallel technical advances in DNA enrichment (e.g., hybridization capture) and

high-throughput next-generation sequencing (NGS) techniques, now allow ancient nuclear genomes to be sequenced to high coverage and accuracy (e.g., Rasmussen et al. 2010; Meyer et al. 2012; Prüfer et al. 2014). Indeed, the field of ancient DNA was rapidly transformed by the application of next generation sequencing technologies (Margulies et al. 2005) to ancient DNA (Poinar et al. 2006). Since then, partial or complete nuclear genomes of 4,000-60,000 year old woolly mammoths (Miller et al. 2008; Lynch et al. 2015; Palkopoulou et al. 2015), three 75,000-130,000 year old extinct camelids (Heintzman et al. 2015), three 16,000-~700,000 year old horses (Orlando et al. 2013; Schubert et al. 2014), a 4,500 year-old Saqqaq paleoeskimo (Rasmussen et al. 2010), several Neanderthals (Green et al. 2010; Prüfer et al. 2014), a previously unknown human lineage (Denisovans) from Siberia (Reich et al. 2010; Meyer et al. 2012), and numerous anatomically modern humans from the Late Pleistocene (e.g., Fu et al., 2014; Rasmussen et al. 2014) have been published. Unfortunately, this tour de force technique has so far been of very limited use to study the physiological peculiarities of extinct species as, due to the random nature by which these data are obtained, some sections of the genome are sequenced several times while others are missed altogether. For instance, about 50% of all DNA sequences—2.1 billion base pairs—in the original 0.8× coverage mammoth genome were not sequenced at all (Miller et al. 2008). Moreover, low coverage genomes contain too many errors due to post-mortem damage of the ancient DNA templates and errors introduced during the sequencing process to be able to offer insights into the physiology of the individual studied. The only way to trust ancient gene sequences obtained with the shotgun approach is to sequence so heavily that most regions are covered many times over, as was done for the two recent mammoth studies (Lynch et al. 2015; Palkopoulou et al. 2015). Given the high levels of post-mortem, exogenous DNA in most ancient specimens (>50 to >95%; e.g., Poinar et al. 2006; Green et al. 2010), this approach quickly becomes cost prohibitive, especially when one considers that studies should target multiple individuals of a species, as intraspecific variation often exists in such traits.

Fortunately, a recently developed approach called hybridization capture enables the targeting and enrichment of millions of base pairs within the genome (Hodges et al. 2007), even from degraded DNA with heavy exogenous contamination (Burbano et al. 2010; Skoglund et al. 2014), something that could never be achieved using PCR. Remarkably, this approach has also allowed for the retrieval of complete mitochondrial genomes from >300,000 year old non-permafrost preserved human and cave bear (*Ursus spelaeus* (Rosenmüller and Heinroth, 1794)) samples (Dabney et al. 2013; Meyer et al. 2013). The lower stringency of this approach moreover permits capture of gene coding sequences from lineages that diverged hundreds of millions of years ago (Li et al. 2013), with the caveat that capture efficiency is inversely proportional to genetic distance. Nonetheless, in combination with next generation sequencing platforms, this approach has recently demonstrated the simultaneous, high coverage retrieval (~99%) of dozens of nuclear genes from five ~1,000 year old Steller's sea cow samples using bait designed from dugong, manatee, and even elephant and hyrax sequence (Mirceta et al. 2013; Springer et al. 2015), the latter two of which diverged from sirenians in the Paleocene (~60 Ma; Meredith et al. 2011; Springer et al. 2015). The validation of cross-ordinal capture from ancient genomes opens the door to studying hundreds of vertebrate (largely avian and mammal) taxa with Pleistocene and Holocene extinction histories (e.g., glyptodonts, ground sloth, thylacine, great auk, elephant bird, among many others), thus providing unparalleled insights into the phylogenetic affinities and divergence times of these species (see Fig. 1). This approach will also permit comprehensive studies on phenotypic evolution that focus, for example, on gene networks underlying tooth- and finger-loss in Steller's sea cows, and the long-hair/short-tail phenotypes of woolly mammoths and woolly rhinos. Given the large number of loci involved in these traits, such studies will benefit from positive selection analyses that help identify genes potentially underlying adaptive changes, as has been conducted for archaic humans (Meyer et al. 2012; Prüfer et al. 2014), mammoths (Lynch et al. 2015), and horses (Orlando et al. 2013;

Schubert et al. 2014). As shown above, however, the gold standard will remain functional verification.

The development of improved sequencing platforms able to sequence billions of DNA fragments in combination with whole genome enrichment and capture—whereby genomes of extant species are fragmented, biotinylated, and used to capture corresponding sequences from closely related extinct taxa (Enk and Poinar 2014)—will further increase the efficiency of ancient DNA retrieval. This relatively low cost, massively parallel targeted sequencing approach also holds the potential for large-scale population-level studies whereby, for example, the complete genomes of Siberian mammoths from relatively warm interglacial periods can be genetically compared and contrasted with individuals living during glacial maxima; direct geographical comparisons among Siberian, Spanish and North American populations may also be possible.

As proteins tend to be more resistant than nucleic acids to degradation, ancient proteomic methodologies promise to substantially extend both the age and range of organisms beyond that currently available to DNA-based approaches, while offering additional benefits such as developmental and tissue specific expression patterns (Cappellini et al. 2014). Specifically, liquid chromatography–tandem mass spectrometry (LC–MS/MS) techniques allow ancient peptides to be directly sequenced from sub-fossil material. To date, this approach has primarily been employed to obtain collagen sequences (which is highly abundant in bone) and elucidate the phylogenetic position of extinct species from temperate and tropical latitudes for which ancient DNA methodologies have failed, such as the endemic South American notoungulates (Welker et al. 2015; Buckley 2015) and the enigmatic ‘Malagasy aardvark’ of Madagascar (Buckley 2013). However, the great potential of LC–MS/MS for paleophysiological characterizations of extinct species was perhaps best demonstrated by the partial retrieval of 126 mostly low-abundance bone matrix and blood plasma proteins (Cappellini et al. 2011) from the same 43,000 year-old woolly mammoth femur sample used in the *MC1R* and hemoglobin studies described above. As generally comparable results have

451 been obtained from archeological bovine samples dating back hundreds of thousands of years
452 (Buckley and Wadsworth, 2014; Wadsworth and Buckley 2014), with a partial MS-generated
453 collagen sequence even reported from 80 million year old hadrosaur (*Brachylophosaurus*
454 *canadensis* (Sternberg, 1953); Schweitzer et al. 2009), this approach currently offers the best
455 hope of directly resurrecting pre-Pleistocene phenotypes.

456 Another recent development, which may have a major impact on the field lies in the ability
457 to reveal epigenetic changes between the genomes of modern and ancient specimens. The first
458 publication that showed that it is in principle possible to study DNA methylation patterns
459 from ancient DNA sequences was published in 2010 (Briggs et al. 2010). Briefly, ancient
460 DNA extracts were first treated with uracil N-glycosylase (UNG), an enzyme that removes
461 uracils, which are common in ancient DNA due to cytosine deamination (Pääbo 1989;
462 Hofreiter et al. 2001); as DNA breaks at these abasic sites, DNA damage artifacts are
463 effectively removed from the resulting sequences. In contrast, methylated cytosins—a
464 hallmark of epigenetic modifications—are deaminated to thymins, which are not removed by
465 UNG and thus appear as C→T changes in the sequences of UNG treated ancient DNA.
466 Alternatively, cytosine methylation can also be measured directly on individual genomic
467 regions (Llamas et al. 2012) or inferred from NGS data if polymerases are used that cannot
468 read over uracils during the initial library amplification (Pedersen et al. 2014). Despite the
469 large potential of ancient epigenetics, there are so far only a handful of studies on this topic
470 (Briggs et al. 2010; Llamas et al. 2012; Pedersen et al. 2014; Gokhman et al. 2014; Smith et
471 al. 2014). While the first three were mostly proof-of-principle publications, Gokhman et al.
472 (2014) found a number of regions that were differentially methylated between modern
473 humans, Denisovans, and Neanderthals. Some of these are quite intriguing as they affect
474 regions that seem to make morphological sense, such as a methylation difference in the
475 HOXD cluster, which affects limb development (Gokhman et al. 2014). However,
476 methylation of the HOXD complex is preserved across great apes raising questions about the
477 reliability of the results (Schneider et al. 2014), although the authors of the original study took

considerable effort to exclude any artefactual results. Even more surprising was the finding of hypermethylation of a non-protein-coding RNA gene (H19) in Neanderthals (Gokhman et al. 2014), as methylation of this locus is conserved across 170 million years of evolution in placental and therian mammals (Schneider et al. 2014), arguing for at least some false positive results in this study. A possible cause may lie in the fact that DNA methylation patterns diminish over time in the fossil record (Smith et al. 2014), although the authors of the Neanderthal/Denisovan study tried to take this factor into account. Thus, while ancient epigenetics may have substantial promise, further studies are required to reveal both its potential and pitfalls (see Orlando et al. 2015 for a comprehensive review on this topic).

Perhaps the most exciting future perspective for the field lies in available technologies that may soon allow ancient DNA adherents to overcome limitations inherent to single gene targeting and/or the use of inappropriate *in vivo/in vitro* genetic background models (e.g., mouse, chicken). For instance, available genome editing techniques that enable precisely targeted modifications (i.e., mutations, insertions, deletions) to be introduced into living cells (Gaj et al. 2013), in conjunction with recent developments in cell biology that allow differentiated adult cells to be reprogrammed into pluripotent stem cells (Takahashi and Yamanaka 2006), theoretically provides a means to simultaneously engineer numerous gene mutations from an extinct species into cells of a closely related taxon. Subsequent *in vitro* differentiation of these cells into isolated tissue cultures of interest (e.g., retinal cells, melanocytes) can then be used to faithfully reconstruct and study complex biological phenotypes of long dead species. Several authors have speculated this approach could one day even allow for the ‘de-extinction’ of iconic Ice Age (and more recently extinct) species (Zimmer, 2013; Seddon et al. 2014), though considerable ethical and technological hurdles remain.

Acknowledgements

KLC acknowledges funding from the University of Manitoba Research Grants Program (URGP), and from the Natural Sciences and Engineering Research Council (NSERC) of Canada Discovery (238838) and Accelerator Supplement Grants (412336) programs. MH is supported by ERC consolidator grant 310763 GeneFlow. We thank A. Signore, M. Gaudry, and two anonymous reviewers for helpful critiques on an earlier draft of this manuscript.

References

- Agarwal, P., Wylie, J.N., Galceran, J., Arkhitko, O., Li, C., Deng, C., Grosschedl, R., and Bruneau, B.G. 2003. Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. *Development*, **130**(3): 623–633. doi:10.1242/dev.00191.
- Boeskorov, G.G., Lazarev, P.A., Sher, A.V., Davydov, S.P, Bakulina, N.T., Shchelchkova, M.V., Binladen, J., Willerslev, E., Buigues, B., and Tikhonov, A.N. 2011. Woolly rhino discovery in the lower Kolyma River. *Quat. Sci. Rev.* **30**(17): 2262–2272. doi:10.1016/j.quascirev.2011.02.010.
- Buckley, M. 2013. A molecular phylogeny of Plesiorcyteropus reassigns the extinct mammalian order ‘Bibymalagasias’ [online]. *PloS ONE*, **8**(3): e59614. doi:10.1371/journal.pone.0059614.
- Buckley, M. 2015 Ancient collagen reveals evolutionary history of the endemic South American ‘ungulates’. *Proc. R. Soc. B* **282**(1806): 20142671. doi:10.1098/rspb.2014.2671.

- 529 Buckley, M., and Wadsworth, C. 2014. Proteome degradation in ancient bone: diagenesis and
530 phylogenetic potential. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **41**: 69–79.
531 doi:10.1016/j.palaeo.2014.06.026.
- 532
- 533 Burbano, H.A., Hodges, E., Green, R.E., Briggs, A.W., Krause, J., Meyer, M., Good, J.M.,
534 Maricic, T., Johnson, P.L.F., Xuan, Z., Rooks, M., Bhattacharjee, A., Brizuela, L., Albert,
535 F.W., de la Rasilla, M., Fortea, J., Rosas, A., Lachmann, M., Hannon G.J., and Pääbo, S.
536 2010. Targeted investigation of the Neandertal genome by array-based sequence capture.
537 *Science*, **328**(5979): 723–725. doi:10.1126/science.1188046.
- 538
- 539 Campbell, K.L., Roberts, J.E.E., Watson, L.N. Stetefeld, J., Sloan, A.M., Signore, A.V.,
540 Howatt, J.W., Tame, J.R.H., Rohland, N., Shen, T.-J., Austin, J.J., Hofreiter, M., Ho, C.,
541 Weber, R.E., and Cooper, A. 2010. Substitutions in woolly mammoth hemoglobin confer
542 biochemical properties adaptive for cold tolerance. *Nat. Genet.* **42**(6): 536–540.
543 doi:10.1038/ng.574.
- 544
- 545 Cappellini, E., Jensen, L.J., Szklarczyk, D., Ginolhac, A., da Fonseca, R.A., Stafford Jr.,
546 T.W., Holen, S.R., Collins, M.J., Orlando, L., Willerslev, E., Gilbert, M.T.P., and Olsen,
547 J.V. (2011). Proteomic analysis of a Pleistocene mammoth femur reveals more than one
548 hundred ancient bone proteins. *J. Proteome Res.* **11**(2): 917–926. doi:10.1021/pr200721u.
- 549
- 550 Cappellini, E., Collins, M.J., and Gilbert, M.T.P. 2014. Unlocking ancient protein
551 palimpsests. *Science*, **343**(6177): 1320–1322. doi:10.1126/science.1249274.
- 552
- 553 Chang, B.S., Jönsson, K., Kazmi, M.A., Donoghue, M.J., and Sakmar, T.P. 2002. Recreating
554 a functional ancestral archosaur visual pigment. *Mol. Biol. Evol.* **19**(9): 1483–1489.
- 555

- 556 Cheng, X., Jin, J., Hu, L., Shen, D., Dong, X., Samie, M.A. Knoff, J., Eisinger, B., Liu, M.-l.,
557 Huang, S.M., Caterina, M.J., Dempsey, P., Michael, L.E., Dlugosz, A.A., Andrews, N.C.,
558 Clapham, D.E., and Xu, H. 2010. TRP channel regulates EGFR signaling in hair
559 morphogenesis and skin barrier formation. *Cell*, **141**(2): 331–343.
560 doi:10.1016/j.cell.2010.03.013.
- 561
- 562 Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera,
563 C., García, N., Pääbo, S., Arsuaga, J.-L., and Meyer, M. 2013. Complete mitochondrial
564 genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA
565 fragments. *Proc. Natl. Acad. Sci. U.S.A.* **110**(39): 15758–15763.
566 doi:10.1073/pnas.1314445110.
- 567
- 568 Darwin, C. 1859. On the origins of species by means of natural selection. John Murray,
569 London.
- 570
- 571 Domning, D.P. 1976. An ecological model for late Tertiary sirenian evolution in the North
572 Pacific Ocean. *Syst. Biol.* **25**(4): 352–362. doi:10.2307/2412510.
- 573
- 574 Enard, W., Przeworski, M., Fisher, S.E., Lai, C.S.L., Wiebe, V., Kitano, T., Monaco, A.P.,
575 and Pääbo, S. 2002. Molecular evolution of *FOXP2*, a gene involved in speech and
576 language. *Nature*, **418**(6900): 869–872. doi:10.1038/nature01025.
- 577
- 578 Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S.M., Blass, T., Somel, M., Brückner,
579 M.K., Schreiweis, C., Winter, C., Sohr, R., Becker, L., Wiebe, V., Nickel, B., Giger, T.,
580 Müller, U., Groszer, M., Adler, T., Aguilar, A., Bolle, I., Calzada-Wack, J., Dalke, C.,
581 Ehrhardt, N., Favor, J., Fuchs, H., Gailus-Durner, V., Hans, W., Hölzlwimmer, G.,
582 Javaheri, A., Kalaydjiev, S., Kallnik, M., Kling, E., Kunder, S., Moßbrugger, I., Naton, B.,

- 583 Racz, I., Rathkolb, B., Rozman, J., Schrewe, A., Busch, D.H., Graw, J., Ivandic, B.,
584 Klingenspor, M., Klopstock, T., Ollert, M., Quintanilla-Martinez, L., Schulz, H., Wolf, E.,
585 Wurst, W., Zimmer, A., Fisher, S.E., Morgenstern, R., Arendt, T., Hrabé de Angelis, M.,
586 Fischer, J., Schwarz, J., and Pääbo, S. 2009. A humanized version of Foxp2 affects cortico-
587 basal ganglia circuits in mice. *Cell*, **137**(5): 961–971. doi:10.1016/j.cell.2009.03.041.
588
- 589 Enk, J.M., Devault, A.M., Kuch, M., Murgha, Y.E., Rouillard, J.M., and Poinar, H.N. 2014.
590 Ancient whole genome enrichment using baits built from modern DNA. *Mol. Biol. Evol.*
591 **31**(5): 1292–1294. doi:10.1093/molbev/msu074.
592
- 593 Fisher, D.C., Tikhonov, A.N., Kosintsev, P.A., Rountrey, A.N., Buigues, B., and van der
594 Plicht, J. 2012. Anatomy, death, and preservation of a woolly mammoth (*Mammuthus*
595 *primigenius*) calf, Yamal Peninsula, northwest Siberia. *Quat. Intern.* **255**: 94–105.
596 doi:10.1016/j.quaint.2011.05.040.
597
- 598 Fu, Q., Li, H., Moorjani, P., Jay, F., Slepchenko, S. M., Bondarev, A. A., Johnson, P.L.F.,
599 Aximu-Petri, A., Prüfer, K., de Filippo, C., Meyer, M., Zwyns, N., Salazar-García, D.C.,
600 Kuzmin, Y.V., Keates, S.G., Kosintsev, P.A., Razhev, D.I., Richards, M.P., Peristov, N.V.,
601 Lachmann, M., Douka, K., Higham, T.F.G., Slatkin, M., Hublin, J.-J., Reich, D., Kelso, J.,
602 Bence Viola, T., and Pääbo, S. 2014. Genome sequence of a 45,000-year-old modern
603 human from western Siberia. *Nature*, **514**(7523): 445–449. doi:10.1038/nature13810.
604
- 605 Gaj, T., Gersbach, C.A., and Barbas III, C. F. 2013. ZFN, TALEN, and CRISPR/Cas-based
606 methods for genome engineering. *Trends Biotechnol.* **31**(7): 397–405.
607 doi:10.1016/j.tibtech.2013.04.004.
608

- Garland, T., and Adolph, S.C. 1994. Why not to do two-species comparative studies: limitations on inferring adaptation. *Physiol. Zool.* **67**(4): 797–828.
- Gaucher, E.A., Govindarajan, S., and Ganesh, O.K. 2008. Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature*, **451**(7179): 704–707. doi:10.1038/nature06510.
- Gokhman, D., Lavi, E., Prüfer, K., Fraga, M.F., Riancho, J.A., Kelso, J., Pääbo, S., Meshorer, E., and Carmel, L. 2014. Reconstructing the DNA methylation maps of the Neandertal and the Denisovan. *Science*, **344**(6183): 523–527. doi:10.1126/science.1250368.
- Grandl, J., Hu, H., Bandell, M., Bursulaya, B., Schmidt, M., Petrus, M., and Patapoutian, A. 2008. Pore region of TRPV3 ion channel is specifically required for heat activation. *Nat. Neurosci.* **11**(9): 1007–1013. doi:10.1038/nn.2169.
- Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Hsi-Yang Fritz, M., Hansen, N.F., Durand, E.Y., Malaspinas, A.-S., Jensen, J.D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H.A., Good, J.M., Schultz, R., Aximu-Petri, A., Butthof, A., Höber, B., Höffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E.S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic, D., Kucan, Ž., Gušić, I., Doronichev, V.B., Golovanova, L.V., Lalueza-Fox, C., de la Rasilla, M., Fortea, J., Rosas, A., Schmitz, R.W., Johnson, F.L.P., Eichler, E.E., Falush, D., Birney, E., Mullikin, J.C., Slatkin, M., Nielsen R., Kelso, J., Lachmann, M., Reich, D., and Pääbo, S. 2010. A draft sequence of the Neandertal genome. *Science*, **328**(5979): 710–722. doi:10.1126/science.1188021.

- 635 Greenwood, A.D., Capelli, C., Possnert, G., and Pääbo, S. 1999. Nuclear DNA sequences
636 from late Pleistocene megafauna. *Mol. Biol. Evol.* **16**(11): 1466–1473.
637
- 638 Heintzman, P.D., Zazula, G.D., Cahill, J.A., Reyes, A.V., MacPhee, R.D., and Shapiro, B.
639 2015. Genomic data from extinct North American *Camelops* revise camel evolutionary
640 history. *Mol. Biol. Evol.* Advance Access published June 2, 2015.
641 doi:10.1093/molbev/msv128.
642
- 643 Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A., and Wilson, A.C. 1984. DNA
644 sequences from the quagga, an extinct member of the horse family. *Nature*, **312**(5991):
645 282–284. doi:10.1038/312282a0.
646
- 647 Hodges, E., Xuan, Z., Baliya, V., Kramer, M., Molla, M.N., Smith, S.W., Middle, C.M.,
648 Rodesch, M.J., Albert, T.J., Hannon, G.J., and McCombie, W.R. 2007. Genome-wide in
649 situ exon capture for selective resequencing. *Nat. Genet.* **39**(12): 1522–1527.
650 doi:10.1038/ng.2007.42.
651
- 652 Hofreiter, M., Serre, D., Poinar, H. N., Kuch, M., and Pääbo, S. 2001. Ancient DNA. *Nat.*
653 *Rev. Genet.* **2**(5): 353–359. doi:10.1038/35072071.
654
- 655 Huynen, L., Suzuki, T., Ogura, T., Watanabe, Y., Millar, C.D., Hofreiter, M., Smith, C.,
656 Mirmoeini, S., and Lambert, D.M. 2014. Reconstruction and *in vivo* analysis of the extinct
657 *tbx5* gene from ancient wingless moa (Aves: Dinornithiformes). *BMC Evol. Biol.* **14**(1):
658 75. doi:10.1186/1471-2148-14-75.
659

- Itakura, K., Hirose, T., Crea, R., Riggs, A.D., Heyneker, H.L., Bolivar, F., and Boyer, H.W. 1977. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science*, **198**(4321): 1056–1063. doi:10.1126/science.412251.
- Jermann, T.M., Opitz, J.G., Stackhouse, J., and Benner, S.A. 1995. Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. *Nature*, **374**(6517): 57–59. doi:10.1038/374057a0.
- Krause, J., Dear, P.H., Pollack, J.L., Slatkin, M., Spriggs, H., Barnes, I., Lister, A.M., Ebersberger, I., Pääbo, S., and Hofreiter, M. 2006. Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. *Nature*, **439**(7077): 724–727. doi:10.1038/nature04432.
- Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R.E., Burbano, H.A., Hublin, J.-J., Hänni, C., Javier Fortea, J., de la Rasilla, M., Bertranpetit, J., Rosas, A., and Pääbo, S. (2007). The derived *FOXP2* variant of modern humans was shared with Neandertals. *Curr. Biol.* **17**(21): 1908–1912. doi:10.1016/j.cub.2007.10.008.
- Krogh, A. 1929. The progress of physiology. *J. Physiol.* **90**(2): 243–251.
- Lalueza-Fox, C., Römpler, H., Caramelli, D., Stäubert, C., Catalano, G., Hughes, D., Rohland, N., Pilli, E., Longo, L., Condemi, S., de la Rasilla, M., Fortea, J., Rosas, A., Stoneking, M., Schöneberg, T., Bertranpetit, J., and Hofreiter, M. 2007. A melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. *Science*, **318**(5855): 1453–1455. doi:10.1126/science.1147417.

- Li, C., Hofreiter, M., Straube, N., Corrigan, S., and Naylor, G.J. 2013. Capturing protein-coding genes across highly divergent species. *BioTechniques*, **54**(6): 321–326. doi:10.2144/000114039.
- Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature*, **362**(6422): 709–715. doi:10.1038/362709a0.
- Lister, A.M., Sher, A.V., van Essen, H., and Wei, G. 2005. The pattern and process of mammoth evolution in Eurasia. *Quat. Inter.* **126–128**: 49–64. doi:10.1016/j.quaint.2004.04.014.
- Llamas, B., Holland, M.L., Chen, K., Cropley, J.E., Cooper, A., and Suter, C.M. 2012. High-resolution analysis of cytosine methylation in ancient DNA [online]. *PLoS ONE*, **7**(1), e30226. doi:10.1371/journal.pone.0030226.
- Lynch, V., Bedoya-Reina, O.C., Ratan, A., Sulak, M., Drautz-Moses, D.I., Perry, G.H., Miller, W., and Schuster, S.C. 2015. Elephantid genomes reveal the molecular bases of Woolly Mammoth adaptations to the arctic [online]. *bioRxiv*, 018366. doi:10.1101/018366.
- Malcolm, B.A., Wilson, K.P., Matthews, B.W., Kirsch, J.F., and Wilson, A.C. 1990. Ancestral lysozymes reconstructed, neutrality tested, and thermostability linked to hydrocarbon packing. *Nature*, **345**(6270): 86–89. doi:10.1038/345086a0.
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J., Chen, Z., Dewell, S.B., de Winter, A., Drake, J., Du, L., Fierro, J.M., Forte, R., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Heen Ho, C., Hutchison, S.K., Irzyk, G.P., Jando, S.C., Alenquer, M.L.I., Jarvie, T.P., Jirage, K.B., Kim,

J.-B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lee, W.L., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Reifler, M., Ronan, M.T., Roth, G.T., Sarkis, G.J., Fredrik Simons, J., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Willoughby, D.A., Yu, P., Begley, R.F., and Rothberg, J.M. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, **437**(7057): 376–380. doi:10.1038/nature03959.

Maricic, T., Günther, V., Georgiev, O., Gehre, S., Čurlin, M., Schreiweis, C., Naumann, R., Burbano, H.A., Meyer, M., Lalueza-Fox, C., de la Rasilla, M., Rosas, A., Gajović, S., Kelso, J., Enard, W., Schaffner, W., and Pääbo, S. 2013. A recent evolutionary change affects a regulatory element in the human *FOXP2* gene. *Mol. Biol. Evol.* **30**(4): 844–852. doi:10.1093/molbev/mss271.

Meredith, R.W., Janečka, J.E., Gatesy, J., Ryder, O.A., Fisher, C.A., Teeling, E.C., Goodbla, A., Eizirik, E., Simão, T.L.L., Stadler, T., Rabosky, D.L., Honeycutt, R.L., Flynn, J.J., Ingram, C.M., Steiner, C., Williams, T.L., Robinson, T.J., Burk-Herrick, A., Westerman, M., Ayoub, N.A., Springer, M.S., and Murphy, W.J. 2011. Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science*, **334**(6055): 521–524. doi:10.1126/science.1211028.

Meredith, R.W., Gatesy, J., Murphy, W.J., Ryder, O.A., and Springer, M.S. 2009. Molecular decay of the tooth gene enamel (ENAM) mirrors the loss of enamel in the fossil record of placental mammals. *PLoS Genet.* **5**(9): e1000634. doi:10.1371/journal.pgen.1000634.

- 738 Meredith, R.W., Gatesy, J., Emerling, C.A., York, V.M., and Springer, M.S. 2013. Rod
739 monochromacy and the coevolution of cetacean retinal opsins. *PLoS Genet.* **9**(4):
740 e1003432. doi:10.1371/journal.pgen.1003432
741
- 742 Meyer, M., Kircher, M., Gansauge, M.T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay,
743 F., Prüfer, K., de Filippo, C., Sudmant, P.H., Alkan, C., Fu, Q., Do, R., Rohland, N.,
744 Tandon, A., Siebauer, M., Green, R.E., Bryc, K., Briggs, A.W., Stenzel, U., Dabney, J.,
745 Shendure, J., Kitzman, J., Hammer, M.F., Shunkov, M.V., Derevianko, A.P., Patterson, N.,
746 Andrés, A.M., Eichler, E.E., Slatkin, M., Reich, D., Kelso, J., and Pääbo, S. 2012. A high-
747 coverage genome sequence from an archaic Denisovan individual. *Science*, **338**(6104):
748 222–226. doi:10.1126/science.1224344.
749
- 750 Miller, W., Drautz, D.I., Ratan, A., Pusey, B., Qi, J., Lesk, A.M., Tomsho, L.P., Packard,
751 M.D., Zhao, F., Sher, A., Tikhonov, A., Raney, B., Patterson, N., Lindblad-Toh, K.,
752 Lander, E.S., Knight, J.R., Irzyk, G.P., Fredrikson, K.M., Harkins, T.T., Sheridan, S.,
753 Pringle, T., and Schuster, S.C. 2008. Sequencing the nuclear genome of the extinct woolly
754 mammoth. *Nature*, **456**(7220): 387–390. doi:10.1038/nature07446.
755
- 756 Moqrich, A., Hwang, S.W., Earley, T.J., Petrus, M.J., Murray, A.N., Spencer, K.S., Andahazy,
757 M., Story, G.M., and Patapoutian, A. 2005. Impaired thermosensation in mice lacking
758 TRPV3, a heat and camphor sensor in the skin. *Science*, **307**(5714): 1468–1472.
759 doi:10.1126/science.1108609.
760
- 761 Ng, J.K., Kawakami, Y., Buscher, D., Raya, A., Itoh, T., Koth, C.M., Rodriguez Esteban, C.,
762 Rodriguez-Leon, J., Garrity, D.M., Fishman, M.C., and Izpisua Belmonte, J.C. 2002. The
763 limb identity gene *Tbx5* promotes limb initiation by interacting with *Wnt2b* and *Fgf10*.
764 *Development*, **129**(22): 5161–5170.

765

766 Noguchi, H., Campbell, K.L., Ho, C., Unzai, S., Park, S.-Y., and Tame, J.R.H. 2012.

767 Structures of haemoglobin from woolly mammoth in liganded and unliganded states. *Acta*768 *Cryst. D* **68**(11): 1441–1449. doi:10.1107/S0907444912029459.

769

770 Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M.,

771 Cappellini, E., Petersen, B., Moltke, I., Johnson, P.L.F., Fumagalli, M., Vilstrup, J.T.,

772 Raghavan, M., Korneliussen, T., Malaspinas, A.-S., Vogt, J., Szklarczyk, D., Kelstrup,

773 C.D., Vinther, J., Dolocan, A., Stenderup, J., Velazquez, A.M.V., Cahill, J., Rasmussen,

774 M., Wang, X., Min, J., Zazula, G.D., Seguin-Orlando, A., Mortensen, C., Magnussen, K.,

775 Thompson, J.F., Weinstock, J., Gregersen, K., Røed, K.H., Eisenmann, V., Rubin, C.J.,

776 Miller, D.C., Antczak, D.F., Bertelsen, M.F., Brunak, S., Al-Rasheid, K.A.S., Ryder, O.,

777 Andersson, L., Mundy, J., Krogh, A., Gilbert, M.T.P., Kjær, K., Sicheritz-Ponten, T., Juhl

778 Jensen, L., Olsen, J.V., Hofreiter, M., Nielsen, R., Shapiro, B., Wang, J., and Willerslev, E.

779 2013. Recalibrating *Equus* evolution using the genome sequence of an early Middle780 Pleistocene horse. *Nature*, **499**(7456): 74–78. doi:10.1038/nature12323.

781

782 Orlando, L., Gilbert, M. T. P., and Willerslev, E. 2015. Reconstructing ancient genomes and

783 epigenomes. *Nat. Rev. Genet.* Advance Access published June 9, 2015.

784 doi:10.1038/nrg3935.

785

786 Pääbo, S. 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic

787 amplification. *Proc. Natl. Acad. Sci. U.S.A.* **86**(6): 1939–1943.

788

789 Pask, A.J., Behringer, R.R., and Renfree, M.B. 2008. Resurrection of DNA function *in vivo*790 from an extinct genome [online]. *PLoS ONE*, **3**(5): e2240.

791 doi:10.1371/journal.pone.0002240.

792

793 Palkopoulou, E., Mallick, S., Skoglund, P., Enk, J., Rohland, N., Li, H., Omrak, A.,
794 Vartanyan, S., Poinar, H., Götherström, A., Reich, D., and Dalén, L. 2015. Complete
795 genomes reveal signatures of demographic and genetic declines in the woolly mammoth.
796 *Curr. Biol.* **25**(10): 1395–1400. doi:10.1016/j.cub.2015.04.007.

797

798 Pauling, L., and Zuckerkandl, E. 1963. Chemical paleogenetics. *Acta Chem. Scand.* **17**: S9–
799 S16. doi:10.3891/acta.chem.scand.17s-0009.

800

801 Pedersen, J.S., Valen, E., Velazquez, A.M.V., Parker, B.J., Rasmussen, M., Lindgreen, S.,
802 Lilje, B., Tobin, D.J., Kelly, T.K., Vang, S., Andersson, R., Jones, P.A., Hoover, C.A.,
803 Tikhonov, A., Prokhortchouk E., Rubin, E.M., Sandelin, A., Gilbert, M.T.P., Krogh, A.,
804 Willerslev, E., and Orlando, L. 2014. Genome-wide nucleosome map and cytosine
805 methylation levels of an ancient human genome. *Genome Res.* **24**(3): 454–466. doi:
806 10.1101/gr.163592.113.

807

808 Phillips, M.J., Gibb, G.C., Crimp, E.A., and Penny, D. 2010. Tinamous and moa flock
809 together: mitochondrial genome sequence analysis reveals independent losses of flight
810 among ratites. *Syst. Biol.* **59**(1): 90–107. doi:10.1093/sysbio/syp079.

811

812 Poinar, H.N., Schwarz, C., Qi, J., Shapiro, B., MacPhee, R.D., Buigues, B., Tikhonov, A.,
813 Huson, D.H., Tomsho, L.P., Auch, A., Rampp, M., Miller, W., and Schuster, S.C. 2006.
814 Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science*,
815 **311**(5759): 392–394. doi:10.1126/science.1123360.

816

817 Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A.,
818 Renaud, G., Sudmant, P.H., de Filippo, C., Li, H., Mallick, W., Dannemann, M., Fu, Q.,

Kircher, M., Kuhlwilm, M., Lachmann, M., Meyer, M., Ongyerth, M., Siebauer, M.,
Theunert, C., Tandon, A., Moorjani, P., Pickrell, J., Mullikin, J.C., Vohr, S.H., Green,
R.E., Hellmann, I., Johnson, P.L.F., Blanche, H., Cann, H., Kitzman, J.O., Shendure, J.,
Eichler, E.E., Lein, E., Bakken, T.E., Golovanova, L.V., Doronichev, V.B., Shunkov,
M.V., Derevianko, A.P., Viola, B., Slatkin, M., Reich, D., Kelso, J., and Pääbo, S. 2014.
The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*,
505(7481): 43–49. doi:10.1038/nature12886.

Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, I., Metspalu,
M., Metspalu, E., Kivisild, T., Gupta, R., Bertalan, M., Nielsen, K., Gilbert, M.T.P., Wang,
Y., Raghavan, M., Campos, P.F., Munkholm Kamp, H., Wilson, A.S., Gledhill, A.,
Tridico, S., Bunce, M., Lorenzen, E.D., Binladen, J., Guo, X., Zhao, J., Zhang, X., Zhang,
H., Li, Z., Chen, M., Orlando, L., Kristiansen, K., Bak, M., Tommerup, N., Bendixen, C.,
Pierre, T.L., Grønnow, B., Meldgaard, M., Andreassen, C., Fedorova, S.A., Osipova, L.P.,
Higham, T.F.G., Bronk Ramsey, C., Hansen, T.V.O., Nielsen, F.C., Crawford, M.H.,
Brunak, S., Sicheritz-Pontén, T., Villems, R., Nielsen, R., Krogh, A., Wang, J., and
Willerslev, E. 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo.
Nature, **463**(7282): 757–762. doi:10.1038/nature08835.

Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford Jr, T.W.,
Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S.M., Poznik, G.D., Gudmundsdottir,
V., Yadav, R., Malaspinas, A.-S., Stockton White V, S., Allentoft, M.E., Cornejo, O.E.,
Tambets, K., Eriksson, A., Heintzman, P.D., Karmin, M., Sand Korneliussen, T., Meltzer,
D.J., Pierre, T.L., Stenderup, J., Saag, L., Warmuth, V.M., Lopes, M.C., Malhi, R.S.,
Brunak, S., Sicheritz-Pontén, T., Barnes, I., Collins, M., Orlando, L., Balloux, F., Manica,
A., Gupta, R., Metspalu, M., Bustamante, C.D., Jakobsson, M., Nielsen, R., and

Willerslev, E. 2014. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature*, **506**(7487): 225–229. doi:10.1038/nature13025.

Rees, J.L. 2003. Genetics of hair and skin color. *Ann. Rev. Genet.* **37**(1): 67–90. doi:10.1146/annurev.genet.37.110801.143233.

Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L.F., Maricic, T., Good, J.M., Marques-Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E.E., Stoneking, M., Richards, M., Talamo, S., Shunkov, M.V., Derevianko, A.P., Hublin, J.-J., Kelso, J., Slatkin, M., and Pääbo, S. 2010. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*, **468**(7327): 1053–1060. doi:10.1038/nature09710.

Roca, A.L., Ishida, Y., Nikolaidis, N., Kolokotronis, S.-O., Fratpietro, S., Stewardson, K., Hensley, S., Tisdale, M., Boeskorov, G., and Greenwood, A.D. 2009. Genetic variation at hair length candidate genes in elephants and the extinct woolly mammoth [online]. *BMC Evol. Biol.* **9**(1): 232. doi:10.1186/1471-2148-9-232.

Rohland, N., Malaspinas, A.-S., Pollack, J.L., Slatkin, M., Matheus, P., and Hofreiter, M. 2007. Proboscidean mitogenomics: chronology and mode of elephant evolution using mastodon as outgroup [online]. *PLoS Biol.* **5**(8): e207. doi:10.1371/journal.pbio.0050207

Rohland, N., Reich, D., Mallick, S., Meyer, M., Green, R.E., Georgiadis, N.J., Roca, A.L., and Hofreiter, M. 2010. Genomic DNA sequences from mastodon and woolly mammoth reveal deep speciation of forest and savanna elephants [online]. *PLoS Biol.* **8**(12): e1000564. doi:10.1371/journal.pbio.1000564.

- 872 Römpler, H., Rohland, N., Lalueza-Fox, C., Willerslev, E., Kuznetsova, T., Rabeder, G.,
873 Bertranpetit, J., Schöneberg, T., and Hofreiter, M. 2006. Nuclear gene indicates coat-color
874 polymorphism in mammoths. *Science*, **313**(5783): 62. doi:10.1126/science.1128994.
875
- 876 Schneider, E., El Hajj, N., and Haaf, T. 2014. Epigenetic information from ancient DNA
877 provides new insights into human evolution. *Brain Behav. Evol.* **84**(3): 169–171.
878 doi:10.1159/000365650.
879
- 880 Schubert, M., Jónsson, H., Chang, D., Der Sarkissian, C., Ermini, L., Ginolhac, A.,
881 Albrechtsen, A., Dupanloup, I., Foucal, A., Petersen, B., Fumagalli, M., Raghavan, M.,
882 Seguin-Orlando, A., Korneliussen, T.S., Velazquez, A.M.V., Stenderup, J., Hoover, C.A.,
883 Rubin, C.-J., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., MacHugh, D.E.,
884 Kalbfleisch, T., MacLeod, J.N., Rubin, E.M., Sicheritz-Ponten, T., Andersson, L.,
885 Hofreiter, M., Marques-Bonet, T., Gilbert, M.T.P., Rasmus Nielsen, R., Excoffier, L.,
886 Willerslev, E., Shapiro, B., and Orlando, L. 2014. Prehistoric genomes reveal the genetic
887 foundation and cost of horse domestication. *Proc. Natl. Acad. Sci. U.S.A.* **111**(52): E5661-
888 E5669. doi: 10.1073/pnas.1416991111.
889
- 890 Schweitzer, M.H., Zheng, W., Organ, C.L., Avci, R., Suo, Z., Freimark, L.M., Lebleu, V.S.,
891 Duncan, M.B., Heiden, M.G.V., Neveu, J.M., Lane, W.S., Cottrell, J.S., Horner, J.R.,
892 Cantley, L.C., Kalluri, R., and Asara, J.M. 2009. Biomolecular characterization and protein
893 sequences of the Campanian hadrosaur *B. canadensis*. *Science*, **324**(5927): 626–631.
894 doi:10.1126/science.1165069.
895
- 896 Seddon, P.J., Moehrensclager, A., and Ewen, J. 2014. Reintroducing resurrected species:
897 selecting DeExtinction candidates. *Trends Ecol. Evol.* **29**(3): 140–147.
898 doi:10.1016/j.tree.2014.01.007.

899

900 Shapiro, B., and Hofreiter, M. 2014. A paleogenomic perspective on evolution and gene
901 function: new insights from ancient DNA. *Science*, **343**(6169): 1236573.
902 doi:10.1126/science.1236573.

903

904 Shi, Y., and Yokoyama, S. 2003. Molecular analysis of the evolutionary significance of
905 ultraviolet vision in vertebrates. *Proc. Natl. Acad. Sci. U.S.A.* **100**(14): 8308–8313.
906 doi:10.1073/pnas.1532535100.

907

908 Skoglund, P., Northoff, B.H., Shunkov, M.V., Derevianko, A.P., Pääbo, S., Krause, J., and
909 Jakobsson, M. 2014. Separating endogenous ancient DNA from modern day contamination
910 in a Siberian Neandertal. *Proc. Natl. Acad. Sci. U.S.A.* **111**(6): 2229–2234.
911 doi:10.1073/pnas.1318934111.

912

913 Smith, M. 1982. Site-directed mutagenesis. *Trends Biochem. Sci.* **7**(12): 440–442.
914 doi:10.1016/S0968-0004(82)80013-3.

915

916 Smith, O., Clapham, A.J., Rose, P., Liu, Y., Wang, J., and Allaby, R.G. 2014. Genomic
917 methylation patterns in archaeological barley show de-methylation as a time-dependent
918 diagenetic process [online]. *Sci. Rep.* **4**: 5559. doi:10.1038/srep05559.

919

920 Springer, M.S., Signore, A.V., Paijmans, J.L.A., Domning, D.P., Velez-Juarbe, J., Bauer, C.,
921 He, K., Crerar, L., Campos, P.F., Murphy, W.J., Meredith, R.M., Gatesy, J., Willerslev, E.,
922 MacPhee, R.D.E., Hofreiter, M., and Campbell, K.L. 2015. Interordinal gene capture, the
923 phylogenetic position of Steller's sea cow based on molecular and morphological data, and
924 the macroevolutionary history of Sirenia. *Mol. Phylo. Evol.* **91**(10): 178–
925 193. doi:10.1016/j.ympev.2015.05.022.

926

927 Stackhouse, J., Presnell, S.R., McGeehan, G.M., Nambiar, K.P., and Benner, S.A. 1990. The
928 ribonuclease from an extinct bovid ruminant. *FEBS Lett.* **262**(1): 104–106.
929 doi:10.1016/0014-5793(90)80164-E.

930

931 Stuart, A.J. 1991. Mammalian extinctions in the Late Pleistocene of northern Eurasia and
932 North America. *Biol. Rev. Camb. Philos. Soc.* **66**(4): 453–562. doi:10.1111/j.1469-
933 185X.1991.tb01149.x.

934

935 Takahashi, K., and Yamanaka, S. 2006. Induction of pluripotent stem cells from mouse
936 embryonic and adult fibroblast cultures by defined factors. *Cell*, **126**(4): 663–676.
937 doi:10.1016/j.cell.2006.07.024.

938

939 Thornton, J.W., Need, E., and Crews, D. 2003. Resurrecting the ancestral steroid receptor:
940 ancient origin of estrogen signaling. *Science*, **301**(5640): 1714–1717.
941 doi:10.1126/science.1086185.

942

943 Thornton, J.W. 2004. Resurrecting ancient genes: experimental analysis of extinct molecules.
944 *Nat. Rev. Genet.* **5**(5): 366–375. doi:10.1038/nrg1324.

945

946 Tridico, S.R., Rigby, P., Kirkbride, K.P., Haile, J., and Bunce, M. 2014. Megafaunal split
947 ends: microscopical characterisation of hair structure and function in extinct woolly
948 mammoth and woolly rhino. *Quat. Sci. Rev.* **83**: 68–75.
949 doi:10.1016/j.quascirev.2013.10.032.

950

- Weber, R.E., and Campbell, K.L. 2011. Temperature dependence of haemoglobin-oxygen affinity in heterothermic vertebrates: mechanisms and biological significance. *Acta Physiol.* 202(3): 549–562. doi:10.1111/j.1748-1716.2010.02204.x.
- Weber, R.E., Fago, A., and Campbell, K.L. 2014. Enthalpic partitioning of the reduced temperature sensitivity of O₂ binding in bovine hemoglobin. *Comp. Biochem. Physiol. A.* 176(10): 20–25. doi:10.1016/j.cbpa.2014.06.012.
- Welker, F., Collins, M.J., Thomas, J.A., Wadsley, M., Brace, S., Cappellini, E., Turvey, S.T., Reguero, M., Gelfo, J.N., Kramarz, A., Burger, J., Thomas-Oates, J., Ashford, D.A., Ashton, P.D., Rowsell, K., Porter, D.M., Kessler, B., Fischer, R., Baessman, C., Kaspar, S., Olsen, J.V., Kiley, P., Elliott, J.A., Kelstrup, C.D., Mullin, V., Hofreiter, M., Willerslev, E., Hublin, J.-J., Orlando, L., Barnes, I., and MacPhee, R.D. 2015. Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature*, 522(7554): 81–84. doi:10.1038/nature14249.
- Willerslev, E., Gilbert, M.T.P., Binladen, J., Ho, S.Y., Campos, P.F., Ratan, A., Tomsho, L.P., da Fonseca, R.R., Sher, A., Kuznetsova, T.V., Nowak-Kemp, M., Roth, T.L., Miller, W., and Schuster, S.C. 2009. Analysis of complete mitochondrial genomes from extinct and extant rhinoceroses reveals lack of phylogenetic resolution [online]. *BMC Evol. Biol.* 9(1): 95. doi:10.1186/1471-2148-9-95.
- Workman, C., Dalen, L., Vartanyan, S., Shapiro, B., Kosintsev, P., Sher, A., Gotherstrom, A., and Barnes, I. 2011. Population-level genotyping of coat colour polymorphism in woolly mammoth (*Mammuthus primigenius*). *Quat. Sci. Rev.* 30(17–18): 2304–2308. doi:10.1016/j.quascirev.2010.08.020.

- Wray, G.A. 2007. The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.* **8**(3): 206–216. doi:10.1038/nrg2063.
- Yuan, Y., Shen, T.-J., Gupta, P., Ho, N.T., Simplaceanu, V., Tam, T.C.S., Hofreiter, M., Cooper, A., Campbell, K.L., and Ho, C. 2011. A biochemical-biophysical study of hemoglobins from woolly mammoth, Asian elephant, and humans. *Biochemistry*, **50**(34): 7350–7360. doi:10.1021/bi200777j.
- Yuan, Y., Byrd, C., Shen, T.J., Simplaceanu, V., Tam, T.C.S., and Ho, C. 2013. Role of $\beta/\delta 101\text{Gln}$ in regulating the effect of temperature and allosteric effectors on oxygen affinity in woolly mammoth hemoglobin. *Biochemistry*, **52**(49): 8888–8897. doi:10.1021/bi401087d.
- Zhang, J. 2006. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. *Nat. Genet.* **38**(7): 819–823. doi:10.1038/ng1812.
- Zimmer, C. 2013. Bringing them back to life. *Nat. Geogr. Mag.* **223**(4): 28–41.

Figure legends

Figure 1. Schematic diagram outlining the steps from subfossil material to paleophysiology. Briefly, ancient DNA extracts from bone (or other material) are purified and ligated to adapter strands housing a specimen-specific identifying index sequence. This step allows multiple samples to be pooled for subsequent hybridization capture and/or sequencing (see below); these barcoded libraries moreover can be safely used in any standard DNA lab—and re-amplified if stocks become depleted—as any introduced contamination (i.e., DNA strands lacking an index sequence) is automatically removed from analyses following sequencing. Targeted regions of the genome are enriched via DNA hybridization with biotinylated oligos (typically 60–120 base pairs in length) carrying complementary sequence, captured with magnetic beads, and eluted for subsequent next generation sequencing; non-target DNA is discarded. Retrieved sequences are then assembled or mapped to reference sequences from closely related species. Depending on the specific question, obtained gene/promoter/enhancer sequences can be ectopically expressed for ensuing functional/structural analyses.

1011 Table 1. Extinct species as exemplars of the August Krogh Principle.

1012

Ideally suited to test	Species	1013
Cold tolerance	Woolly mammoth, woolly rhinoceros, Steller's sea cow	
Rapid dietary change	Cave bears	
Fully aquatic adaptation	Great auk	
Gigantism	Haast's eagle, Steller's sea cow, giant sloth	
Wing reduction	Moas, elephant bird	
Hair and tail length	Woolly mammoth, woolly rhinoceros	
Finger and tooth loss	Steller's sea cow

Figure 1. Schematic diagram outlining the steps from subfossil material to paleophysiology. Briefly, ancient DNA extracts from bone (or other material) are purified and ligated to adapter strands housing a specimen-specific identifying index sequence. This step allows multiple samples to be pooled for subsequent hybridization capture and/or sequencing (see below); these barcoded libraries moreover can be safely used in any standard DNA lab—and re-amplified if stocks become depleted—as any introduced contamination (i.e., DNA strands lacking an index sequence) is automatically removed from analyses following sequencing. Targeted regions of the genome are enriched via DNA hybridization with biotinylated oligos (typically 60–120 base pairs in length) carrying complementary sequence, captured with magnetic beads, and eluted for subsequent next generation sequencing; non-target DNA is discarded. Retrieved sequences are then assembled or mapped to reference sequences from closely related species. Depending on the specific question, obtained gene/promoter/enhancer sequences can be ectopically expressed for ensuing functional/structural analyses.

182x237mm (300 x 300 DPI)

Draft