Michael Hofreiter | Johanna L. A. Paijmans | Helen Goodchild | Camilla F. Speller | Axel Barlow | Gloria G. Fortes | Jessica A. Thomas | Arne Ludwig | Matthew J. Collins

The future of ancient DNA

Technical advances and conceptual shifts

Suggested citation referring to the original publication:
BioEssays 37 (2015) 3, 284-293
DOI https://doi.org/10.1002/bies.201400160
ISSN (print) 0265-9247
ISSN (online) 1521-1878

The future of ancient DNA: Technical advances and conceptual shifts

Michael Hofreiter[1)2)], Johanna L. A. Paijmans[1)], Helen Goodchild[3)], Camilla F. Speller[3)], Axel Barlow[1)], Gloria G. Fortes[1)], Jessica A. Thomas[2)], Arne Ludwig[4)] and Matthew J. Collins[3)]*

Technological innovations such as next generation sequencing and DNA hybridisation enrichment have resulted in multi-fold increases in both the quantity of ancient DNA sequence data and the time depth for DNA retrieval. To date, over 30 ancient genomes have been sequenced, moving from 0.7× coverage (mammoth) in 2008 to more than 50× coverage (Neanderthal) in 2014. Studies of rapid evolutionary changes, such as the evolution and spread of pathogens and the genetic responses of hosts, or the genetics of domestication and climatic adaptation, are developing swiftly and the importance of palaeogenomics for investigating evolutionary processes during the last million years is likely to increase considerably. However, these new datasets require new methods of data processing and analysis, as well as conceptual changes in interpreting the results. In this review we highlight important areas of future technical and conceptual progress and discuss research topics in the rapidly growing field of palaeogenomics.

Keywords:
ancient DNA; hybridisation capture; multi-locus data; next generation sequencing (NGS); palaeogenomics; population genomics

Introduction

The year 2014 marked the 30th anniversary of ancient DNA (aDNA) research, and it is thus a good time both to highlight the striking transformations the field has undergone over the past 30 years, and to look forward to the future. The first studies on aDNA and RNA – a topic that is often ignored – generally concentrated on simply retrieving and characterising the length of ancient nucleic acid sequences [1]. During the pre-PCR era, only two studies reported aDNA sequences: the seminal study by Higuchi et al. [2] who reported a total of 229 bp of mitochondrial (mt) DNA from the quagga (*Equus quagga quagga*, an extinct, stripeless sub-species of the plains zebra) and partial sequencing of a 3.4 kb fragment cloned from DNA obtained from an Egyptian human mummy [3]. While the first study has withstood the test of time [4], the second is now accepted to have been the result of contamination with modern DNA [5]. The problem of contamination has stayed with aDNA research ever since, and is an issue that has been reviewed multiple times [6, 7].

With the introduction of the polymerase chain reaction (PCR), it became possible to target and replicate specific DNA sequences. The sequences that can be obtained from aDNA are short (apart from exceptionally preserved permafrost samples that allow up to 1.6 kb of DNA to be amplified in one fragment [8]), but by using overlapping PCR fragments it is possible to assemble fairly large DNA sequences, including full mitochondrial genomes or individual nuclear genes. Ancient DNA sequences obtained via PCR have allowed

DOI 10.1002/bies.201400160

[1)] Institute for Biochemistry and Biology, University of Potsdam, Potsdam, Germany
[2)] Department of Biology, University of York, York, UK
[3)] Department of Archaeology, University of York, York, UK
[4)] Department of Evolutionary Genetics, Institute for Zoo and Wildlife Research, Berlin, Germany

***Corresponding author:**
Michael Hofreiter
E-mail: michael.hofreiter@uni-potsdam.de

Abbreviations:
aDNA, ancient DNA; **mtDNA,** mitochondrial DNA; **NGS,** next generation sequencing; **PCA,** principle component analysis; **SNP,** single nucleotide polymorphism.

comprehensive phylogenetic studies and population genetics of both extinct and extant species, as well as functional studies of extinct genes and various aspects of domestication (for reviews, see [6, 9–11]). However, in some fields PCR-based studies remained controversial, such as aDNA studies on human pathogens [12, 13]. Furthermore, complete nuclear palaeogenomes were out of reach for PCR-based approaches.

Almost all of these limitations have been overcome with the development of so-called next generation sequencing (NGS) technologies that allow billions of sequences to be analysed simultaneously [14]. These technologies have not only enabled high-coverage nuclear and mitochondrial genomes to be obtained from aDNA; they have also paved the way for other technical applications that have revolutionised research on aDNA in recent years, such as DNA hybridisation enrichment [15]. DNA hybridisation enrichment not only allows targeting DNA regions up to many megabases in length, but it is also much better suited than PCR for retrieving the short fragments characteristic of aDNA. Hybridisation enrichment has been used to target full mitochondrial genomes [16, 17], as well as for enrichment of nuclear DNA [18, 19]. It has been particularly important in re-establishing studies on ancient pathogen DNA as a viable field of research (see below). Moreover, it has allowed the retrieval of the oldest aDNA sequences to date from non-permafrost samples (~400,000 year old), even from samples for which shotgun sequencing did not reveal any endogenous DNA [20].

NGS has also improved our understanding of the degradation processes that aDNA undergoes over time. Cytosine deamination [21], resulting in C > U changes at regular cytosines and C > T changes at 5-methylated cytosines [22], has allowed increasingly refined distinctions between endogenous and contaminant sequences. It has been shown that cytosine deamination mostly occurs at the ends of aDNA fragments [23, 24] and – albeit with large variations – its frequency increases with the age of a sample [25]. Most interestingly, researchers have started using cytosine deamination to authenticate individual DNA and RNA reads [26–28], as well as selectively targeting and enriching uracil-containing DNA strands [29] – approaches which have enabled the retrieval of partial or complete mitochondrial genomes from a highly contaminated Neanderthal samples [30].

Given all these achievements, the direction and challenges of aDNA research are likely to change radically over the next decades. We will discuss a number of topics in which we foresee future progress in aDNA research. They can be divided into two main categories: first, data generation, which includes technical progress such as lab protocols and bioinformatics; and second, data interpretation, in which we discuss new and revived fields of aDNA research as well as conceptual progress in interpreting the data.

Data generation

Ancient DNA survival is correlated with thermal age

Given the possibilities for aDNA analyses, identifying the right samples for analysis is becoming increasingly important. For bone samples, the amount of amino acid racemisation has been proposed as a reliable proxy in the past [31], but it has been shown that this conclusion was based on misinterpretation of the data [32]. For other samples like sediments or plant tissues, no proxies have been developed at all. However, it has been known for some time that DNA survives longer in cold conditions, ideally frozen, and substantial progress in understanding the degradation process of endogenous DNA has been made [23, 27]. Using well-dated moa bones [33] and coral samples [34], fragmentation rates of both DNA and protein have been directly measured.

To predict aDNA survival globally, we have developed a model for DNA degradation based on environmental temperature. We estimated temperature history (www.thermal-age.eu) from local temperatures using the WorldClim database [35] with corrections for soil depth. Long-term fluctuations in temperature were estimated by calibrating modelled northern hemisphere atmosphere temperatures from [36] against local variation identified in the PMIP2 models (http://pmip2.lsce.ipsl.fr/) of mid-Holocene (6,000 years ago) and Last Glacial Maximum (21,000 years ago) climate. We then applied the model to cases where the DNA degradation rate has been independently measured [37], and found that the model-estimates are close to the empirically observed values. Based on these results, we predicted the survival of DNA fragments of different length across the globe for both open and cave sites (Fig. 1). If the underlying assumption that ancient DNA survival is correlated to temperature is correct, these maps give an impression of the challenges faced by researchers working in low latitudes and on certain continents such as Africa, South America and Australia where poor DNA survival is predicted. They also highlight the potential of caves (and/or deep burial) at low latitudes. Finally, the maps also illustrate the challenges of conventional PCR, which typically targets longer fragments, and the advantage of the short reads that can be retrieved by using NGS with or without hybridisation enrichment.

Optimising extraction methods is crucial for the recovery of aDNA sequence information

Even in a well-preserved sample, aDNA is generally only present in small amounts, and the number of samples available for destructive sampling is often limited. Therefore, maximising DNA extraction efficiency is of vital importance. For bone and dental remains, quite extensive testing and development of extraction methods have been performed [38–40]. The most recent method permits efficient recovery of fragments shorter than 50 bp and has allowed reconstruction of the oldest mitochondrial genomes from non-permafrost samples to date: an ~400,000 year old cave bear and an ~400,000 year old hominin [20, 41, 42]. The importance of short fragment recovery in aDNA research, especially on very old samples, was recently demonstrated with the analysis of an ~700,000-year-old horse sample [43]. The authors of this study calculated that even retrieval of aDNA >1 million years old might be possible, provided very short fragments can be recovered. A comparison of fragment length distributions from the ~700,000 year-old horse and the ~400,000 year old cave bear sequences suggests that substantially more DNA

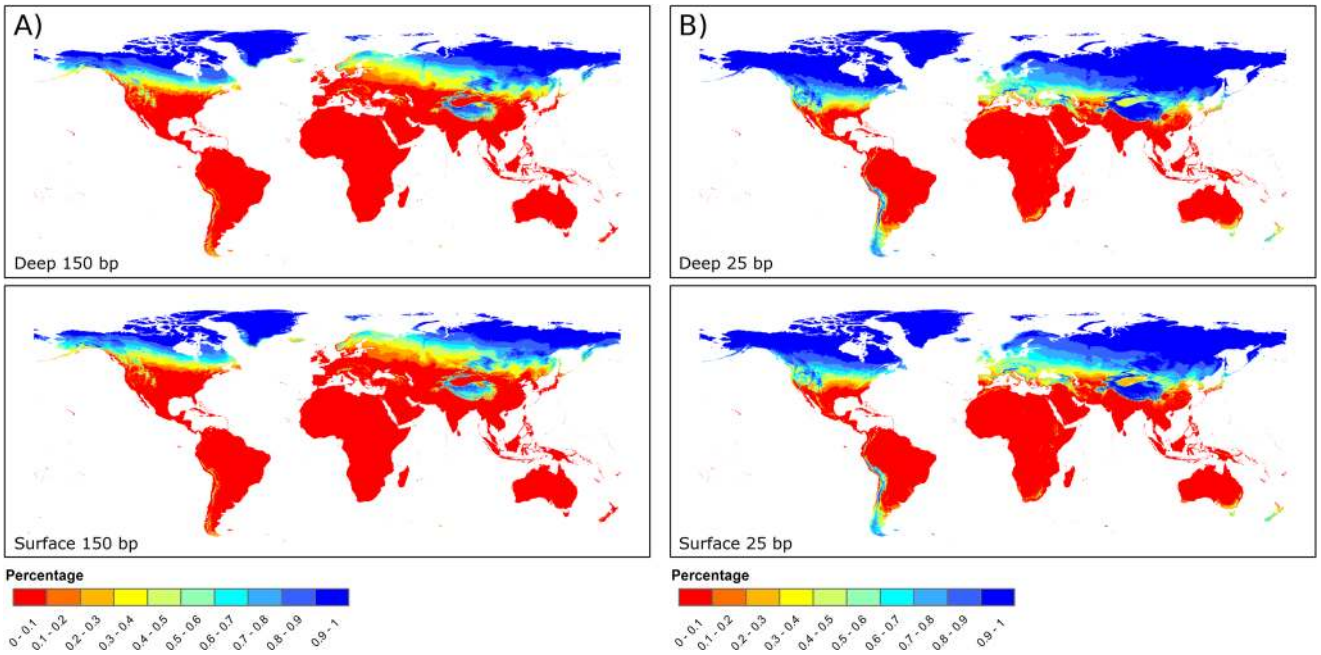

Figure 1. Estimation of DNA survival. Expected survival of DNA after 10,000 years, for 150 bp fragments (**A**) and 25 bp fragments (**B**). While it is clear that the chances of any DNA surviving in desert and tropical regions is minimal for any fragment length, in temperate regions the recovery of short fragments is much more likely than that of longer ones.

may have been recovered from the horse had a more efficient extraction method been used (Fig. 2). Furthermore, there is evidence that – at least in some cases – only a small portion (or even none) of the endogenous DNA in a bone can be solubilised into extraction buffer [44], and even if DNA is obtained, most methods recover only a small proportion of the aDNA present [45]. Improvements in DNA extraction techniques are probably still possible for all ancient substrates; thus, although these experiments are often tedious, we strongly encourage systematic studies aimed at optimising aDNA recovery from all types of substrates in order to increase time depth, breadth of sampling locations and endogenous DNA recovery. If possible, selection of samples with high levels of endogenous DNA also reduces sequencing costs. Hair, although rarely preserved has proved easy to decontaminate [46, 47], while a recent paper highlights the value of using the non-vascularised petrous bone within the skull [48].

Next generation sequencing has radically changed aDNA research

Following DNA extraction, the next step in aDNA analyses is either PCR amplification or, increasingly, the construction of next generation sequencing libraries (a trend that is likely to accelerate). Although the first aDNA studies applying NGS methods used standard library construction protocols, it soon became clear that these result in the loss of a substantial percentage of DNA, and are therefore not really suitable for subfossil samples [49, 50]. However, only a small number of studies have aimed at improving the conversion efficiency of aDNA into NGS libraries [51–53]. For the most commonly used NGS technology in aDNA studies, the Illumina platform, the method currently believed to be the most efficient is a protocol based on single stranded DNA ligation [54] with conversion efficiencies of about 30–70% [55]. Although it is unlikely that values close to 100% are realistic, it might be possible to further improve conversion efficiency.

NGS libraries can be sequenced directly (shotgun metagenomics) or be enriched for certain sequence regions using hybridisation capture and then sequenced. Since introduction of NGS technologies, the most critical parameter for aDNA analyses – the number of reads that can be processed in a single sequencing run – has increased from about 300,000 in the first generation of 454 sequencers [14] to an announced 1.8 billion reads for the next upgrade of the Illumina HiSeq (HiSeqX^{TM}Ten: http://investor.illumina.com/phoenix.zhtml?c=121127&p=irol-newsArticle&ID=1890696). Although there has been a lot of discussion about true single molecule and nanopore sequencing [56, 57], it is uncertain when these technologies may be routinely available for aDNA studies. Independent of the development of new technologies, the current throughput together with foreseeable progress makes whole genome sequencing (of both nuclear and mitochondrial genomes) increasingly cost-effective (Fig. 3). However, unless there is another substantial increase in the number of reads that can be processed, the low percentage of endogenous DNA in many ancient samples makes population genomic projects based on aDNA unfeasible for most species. There may be some exceptions to this prediction though, for example humans and domesticated species for which an increasing number of modern genomes are available for comparison [58] or samples with favourable preservation, such as mammoths with a large fossil record from permafrost regions. For the former species, biological insights are

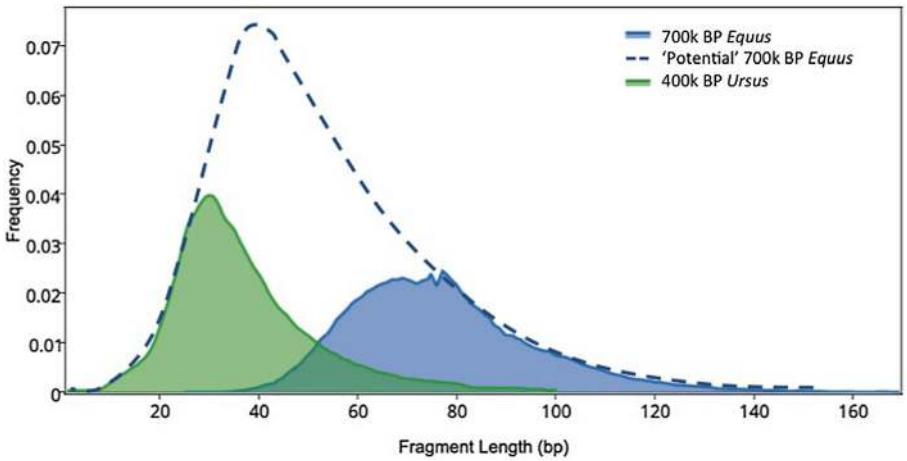

Figure 2. Fragment length distributions. The green graph represents the fragment length distribution from the ~400,000 year old cave bear (*Ursus denigeri*) from Sima de los Huesos [37], where a new extraction method was used, specifically designed for the recovery of short fragments. The blue graph shows the fragment length distribution from the ~700,000 year old horse (*Equus*) from Thistle Creek [49], which was extracted using a standard aDNA extraction method. The dashed line is an extrapolation from the cave bear fragment length distribution, to suggest the potential increase in recovery of short fragments for the horse sample, had the improved extraction method been used.

possible even from individual genomes sequenced to low coverage [53], while for the later species it should be possible to find a sufficient number of samples with high endogenous DNA content.

Hybridisation enrichment as an alternative to shotgun sequencing

For samples with low endogenous DNA content, shotgun sequencing is often financially unfeasible and also inefficient. For such samples, hybridisation enrichment provides an approach to enrich a DNA pool for large genomic regions (for example the exome or mitogenome). Many studies have already used hybridisation enrichment to target ancient mtDNA from a variety of species [16, 20, 59], but there is also a growing number of studies that target chloroplast and nuclear DNA [60, 61] and even whole nuclear genomes [18, 19]. Although the limits of the different hybridisation enrichment techniques for aDNA analysis are still uncertain, many case studies using hybridisation enrichment show that the approach is flexible and sensitive. One study [19] showed that in-solution hybridisation enrichment allowed for the retrieval of over 70% of the original molecules present in the sequencing library, suggesting that hybridisation enrichment can be used with very little loss of endogenous DNA. Unfortunately, enrichment rates for different samples within the same experiment often vary significantly, the reasons for which are often not immediately clear. A recent study [62] compared shotgun data with the data generated after enrichment for the whole genome, and found enrichment ranged from 1.8× up to 14.3×, almost 10-fold different between the individual samples. For mtDNA, the difference between samples is even more pronounced, with enrichment rates of different individuals in one study ranging from 22× to 2,217× [62]. This shows that in order to fully utilise the possibilities of hybridisation enrichment, more studies to understand the effects of different capture parameters are necessary. Therefore, while whole genome capture may one day allow for population-level sampling of ancient genomes, currently the costs for this approach are substantial and the enrichment efficiency not sufficient to allow for large-scale sequencing of ancient genomes for a reasonable fee. In the immediate future, palaeo-population genomics studies (except for unique cases where many samples with high endogenous content are available) are more feasible when combined with hybridisation enrichment for a subset of the genome.

This is exemplified by a recent study [63], which applied hybridisation enrichment to assemble complete exomes for three Neanderthals. Although our understanding of the effect of allelic changes on the phenotype is generally limited without further biochemical experiments, whole exome

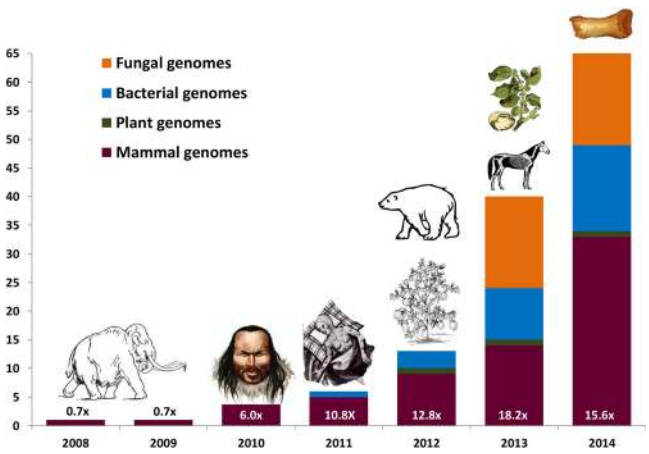

Figure 3. Ancient genomes. Cumulative number of ancient genomes published since the first publication of a draft palaeo-genome in 2008. Numbers in columns indicate cumulative average-fold coverage.

capture of ancient samples provides a promising starting point for future studies into past adaptations. Moreover, the exome is generally more conserved than other regions of the genome, and cross-species capture of exons has already been successfully applied over medium evolutionary distances, even for degraded samples [64, 65]. Thus, exome capture may prove to be more appropriate for the investigation of extinct species where no closely related genome sequence information is available.

An alternative approach is to target specific SNPs that can provide information on a range of biological properties, such as population affinity and ancestry, or can be used to investigate functional or deleterious genes or alleles that may be associated with particular phenotypes [66]. For a number of species, pre-designed SNP arrays are available as ready-to-use assays (e.g. humans [67], domesticates [68], fish [69] and birds [70]), but with the increasing number of reference genomes available, there will be more opportunities for researchers to design project-specific SNP panels aimed to answer their particular research questions. SNP arrays have been sporadically used for ancient samples (e.g. [71]), but this may become more prevalent in future aDNA research; particularly when used as enrichment strategy combined with NGS. However, for the analysis of extinct populations or species, such SNP panels should be used with caution, as they may be affected by ascertainment bias, hence masking any unique variations or overestimate relatedness [72, 73].

Bioinformatic progress lags behind technological progress

Improvements in sequencing technology, extraction, library construction and hybridisation enrichment have resulted in a striking increase in sequence yield, creating new needs and challenges for effective analyses of the sequence reads obtained. Reconstructing the original sequence of a sample can be accomplished either by mapping the reads to a known reference sequence using local alignment algorithms, or by assembling the reads 'de novo' into overlapping scaffolds.

Using a reference sequence for genome reconstruction is most commonly used in aDNA studies as the algorithms are better suited to deal with the low abundance and short lengths of aDNA fragments [74]. However, the algorithms become less effective with increasing sequence divergence between reference and target sequence. This represents an important challenge to be addressed by future studies to ensure aDNA research is not limited to only those taxa where genomic information from closely related modern representatives is available. Improvements in mapping efficiency and accuracy have already been achieved by performing iterative mapping [75, 76]. A combination of de novo contig assembly followed by alignment to related taxa (e.g. [77]) may also provide further enhancements.

BLAST also plays an important role in aDNA bioinformatics, as a tool to detect and filter modern or ancient contaminant sequences. Furthermore, as the fields of ancient substrate metagenomics (e.g. [78]) and contamination detection (e.g. [30]) continue to expand, BLAST and the development of software such as Kraken [79] and One Codex (https://onecodex.com/) will play a key role in identifying and quantifying species of interest within multitaxa pools of sequence reads. It is important to note that taxonomic and genetic identification is entirely dependent on the database of (annotated) sequences used. Accurate identification and information about the sequence query may be biased towards commonly investigated organisms, such as humans, model-organisms and well-characterised pathogens. In the future, BLAST (and other such software) will hopefully provide more accurate annotation, as sequence databases grow, and the corresponding annotation and interpretation improve.

The analytical tools to analyse complete genomes are likewise still under development. Unlike the mitogenome, complete nuclear genomes cannot be projected easily as a single tree. Multivariate analyses such as Principal Component Analysis (PCA) have become a common tool to explore the complex patterns of variation as revealed by multi-locus datasets. However, such analyses can be misleading: first, a PCA compresses the dataset down to basic patterns of variation, which causes a loss of the more subtle (but potentially biologically important) signals; second, the comparative datasets are generally modern-day genomes, which may introduce an ascertainment bias [72, 73]; and third, ancient genomes introduce a temporal aspect that needs to be accounted for in the analysis [80]. New modelling tools such as TreeMix [81] and D-statistics are developed to investigate the genetic affinity of a particular individual compared to other individuals or populations, and help visualise these in a more comprehensive manner. However, the theoretical background, the tools themselves and the interpretation of the results have yet to be standardised.

Data interpretation

New conceptual frameworks for understanding past populations dynamics

The emphasis on mtDNA that has persisted in aDNA research for a long time has almost certainly influenced the way researchers have thought about evolutionary history. In most animal species, mtDNA is inherited through the female lineage without recombination. Mitochondrial DNA sequences can therefore easily be traced to a single most recent common ancestor, and thus conveniently divided into reciprocally monophyletic clades. However, in reality both evolution and human history represent processes during which one state gradually shifts towards another one, without clear boundaries between them. Speciation generally occurs when two populations gradually diverge, with no definable point in history where the populations suddenly become two species. This problem is well reflected by the difficulties researchers have in defining species concepts, and whether particular lineages or clades represent different species, subspecies, populations, or none of these [82, 83]. Similarly, major changes in human history, such as the 'Neolithic revolution', are now universally accepted as gradual and lengthy processes, reflected in the aforementioned case by the adoption of the term 'Neolithic transition'.

Contrary to mtDNA, the nuclear genome represents thousands of recombining loci, under varying selection intensities, and will therefore better reflect the complex and gradual nature of evolution. This is clearly illustrated by the increasing number of studies revealing multiple events of admixture and gene flow, both within modern human populations as well as between the different ancient hominin groups [84, 85]. Thus, rather than representing distinct independent entities, populations should be seen as open pools connected to other pools by larger or smaller rivers. Ancient DNA analyses further complicate, but also enlighten, the picture of past events by adding populations to the analysis that are now extinct: the extinct populations represent pools that are dry today (and therefore invisible), but have been water-filled and connected to other pools in the past. Extant populations are therefore the product of a complex and dynamic history of connections and separations that may not fit the convenient categorical model suggested by mtDNA. Although envisaging the fluidity of population histories as revealed by whole genome analyses is much more difficult (Fig. 4), it is absolutely crucial for understanding evolutionary and historical processes. It should be noted that in this conceptual framework, mtDNA does not become redundant. Rather, information from this and other uniparentally inherited loci (such as the mammalian Y-chromosome) used in conjunction with autosomal data provide the fascinating opportunity to investigate sex-specific processes, such as male- versus female-biased dispersal, gene flow and admixture [86, 87]. For most genome studies, minimal additional efforts are necessary to reconstruct the mitochondrial genome if the nuclear genome is sequenced; thus mitochondrial genomes will (and should) remain a valuable aspect of palaeogenomics.

The sheer time depth that aDNA analyses can reach nowadays (~400,000 years in temperate areas, >700,000 years in permafrost) is difficult to properly comprehend: the development of human civilisation only occupies about 1% of the time scale that can be covered by aDNA. Modelling has shown that all modern humans share a common ancestor as recently as 15,000 years [88], representing just 2% of this time scale. Basically all significant events in human history are squeezed at the end of this time line (Fig. 5). Taking another example, if a population moves only 1 km per year (i.e. about 3 m a day), in 100,000 years it can migrate around the equator about 2½ times. Thus, on the time scale of aDNA studies, there have been countless events of population pools appearing and disappearing, obtaining and losing connections. Therefore, we should expect to see complex series of population replacements, extinctions and gene flow, not only for humans (as reviewed recently in [89]), but for every organism under study.

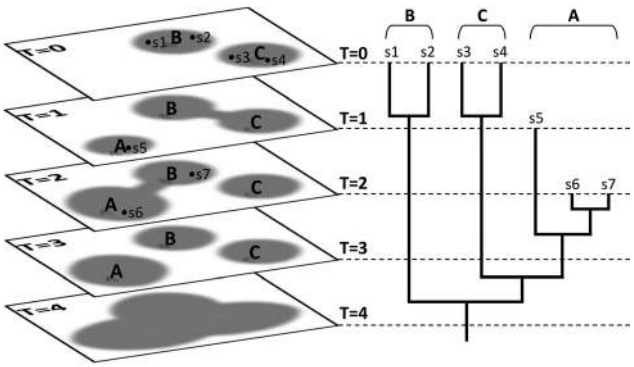

Figure 4. Diagrammatic representation of population histories through the space-time continuum. The rectangular panels on the left show three populations (A, B, C) distributed across a continuous landscape, at different points in time (T = 4–0). White areas in the panels indicate regions of restricted gene flow and grey areas indicate regions of free gene flow. Following the initial divergence of these populations (T = 3), instances of secondary admixture occur between the populations (T = 2 and 1), and one population (A) goes extinct between T = 1 and 0. The right panel shows a hypothetical single locus gene tree generated by sampling seven individuals (s1 to s7) at different time points. With this single genetic locus, we can categorically assign individuals to different clades corresponding to the three populations, but we fail to infer the complex patterns of gene flow that have occurred. Contemporaneous samples (s1-s4) reveal the presence of two populations (B, C), and the addition of ancient samples (s5-s7) reveals the extinct population A. However, we fail to detect secondary gene flow between B and C (T = 1) because the signal of secondary admixture has not been retained in this single locus. Furthermore, the exchange of alleles between A and B (T = 2) leads to the incorrect categorical assignment of individual s7 into population A and potentially an explanation of population replacement rather than admixture. Data from multiple, unlinked genetic markers are required to elucidate the true population histories.

Technological revolution is leading to both novel and resurrected research directions

With the development of new technologies, some fields of aDNA research are being revitalised, and new fields of research are emerging. For example, there has been a renewed focus on the recovery of RNA in desiccated seeds [90], and a complete ancient RNA virus genome was recently sequenced from barley grain over 750 years old [28]. In these materials, RNA may be particularly stable, as shown by the successful germination of a 2,000 year old seed from the Near East [91]. However, the potential for RNA preservation in other substrates has yet to be explored in detail.

Another field of research that has been revitalised by the introduction of NGS and hybridisation enrichment technologies is the study of ancient microbiota. DNA studies of ancient bacteria were highly controversial for many years [12, 13, 92, 93], predominantly due to a lack of independent replication and authentication criteria, which would effectively screen for contamination from ubiquitous environmental microbes. More recent studies have shown that using both shotgun sequencing and hybridisation enrichment approaches it is possible to obtain (more) reliable genetic information from palaeopathogens [94]. To date, palaeopathogen analyses have focused on genetically monomorphic (clonal) pathogens with relatively low mutation rates and 'closed' pan-genomes [95]. For polymorphic microbiota with 'open' pan-genomes and greater rates of horizontal gene transfer (e.g. *Helicobacter pylori*, *Streptococcus agalactiae*, *Escherichia coli*), ancient genome reconstruction will pose more of a challenge, as filtering for contamination from environmental microbiota is difficult [96, 97]. However, as lab protocols and bioinformatics capabilities improve, ancient microbial research will

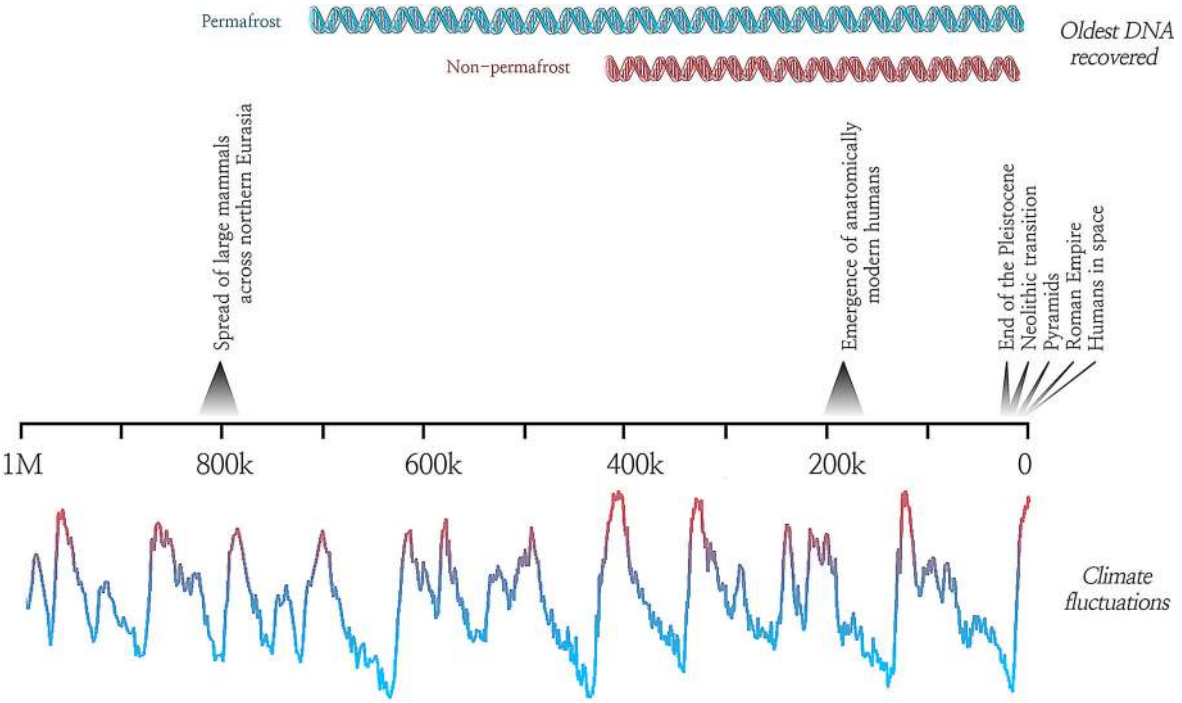

Figure 5. Timescale of ancient DNA analyses. Time line of the last one million years, with some important human events displayed in relation to the oldest aDNA recovered to date from permafrost and non-permafrost samples, respectively. The graph below the timeline reflects global climate fluctuations showing that the timescale of aDNA covers multiple glacial-interglacial cycles.

likely move away from the analysis of single pathogens to population- or community-level analyses, in particular focusing on human microbial communities or 'microbiomes' from a variety of substrates (e.g. dental calculus [94, 98], coprolites [99] and mummified soft tissues [100]).

The rapid progress in aDNA technology provides the opportunity to investigate not only past human health, disease and dietary patterns, but also the co-evolution of humans and pathogens. Modern DNA studies have already detected several human genetic variants which influence susceptibility to common diseases [101]. By sequencing candidate genes, exomes or whole genomes directly from ancient humans and pathogens, it will become possible to directly observe the genetic modifications and adaptations of human populations exposed to pathogens and epidemics, and associate these with specific pathogen mutations or variants through time and space. Thus, the field is likely to move ahead from reporting individual genomes to a population genomics approach in which multiple genomes across space and time are sequenced from individuals and their pathogens, providing a three-dimensional picture of the evolution of pathogens and the human immune system.

The investigation of domestication is also likely to become more multi-faceted in the future, as population-level genome data from ancient samples illuminate this process. The investigation of aDNA has already provided unique glimpses into (human-mediated) population dynamics of companion animals and crop plants at the start of their domestication (e.g. [59, 102]), as well as into functional gene changes associated with domestication (e.g. [103, 104]). By exploring population-level genome data from large numbers of ancient samples, future studies can elucidate with much greater resolution the extent of genetic exchange between ancient domestic and wild populations, as well as allele fluctuations and gene flow between geographical regions [105]. This will provide us with unprecedented insights into processes that drive the continuous change of domestic animals and plants. In addition, it is important to realise that humans have not only mediated genetic changes in the domesticated species, but in turn these species have had an effect on humans as well. A well-known example is the development of lactase persistence in Europe concurrent with the introduction of dairy products [106, 107]. An interesting new research direction in palaeogenomics would be to further investigate how our domesticates have domesticated us.

Moving away from humans and the species directly interacting with them, the new palaeogenomic tools also offer the possibility of investigating the genetic responses of populations of wild species to environmental changes within the last few hundred thousand years. Although understanding the historical processes responsible for shaping patterns of genetic variation is a central aim of phylogeography [108], these studies typically infer changes in allelic diversity and population size indirectly from modern populations. Ancient DNA, in contrast, provides the unique opportunity to sample populations before, during and after periods of environmental change, in order to directly measure shifts in the abundance of alleles. Although this strategy has already been applied to archival samples (for example to demonstrate a reduction in allelic diversity in fish populations following overharvesting [109]), modern palaeogenomic techniques will allow the temporal scale of such studies to be extended considerably. Moreover, such data could allow the investigation of not only

neutral genetic changes but also adaptive ones, providing the unique opportunity to observe Darwinian evolution in real-time even for large species with long generation times, hence fulfilling a long-standing dream of evolutionary biologists. Such studies also provide important information to make more realistic predictions about survival potential of species under future, predicted climatic changes.

Finally, the recent publication of the first ancient nucleosome map and genome-wide survey of cytosine methylation levels gleaned directly from NGS reads indicates the promise of exploring epigenetic information from ancient genomes [110]. Another recent study has characterised the demethylation of cytosines in archaeological plant material over time, most likely due to degradation and subsequent fragmentation of the DNA, which can lead to a loss of methylation signal in unfavourable preservation conditions [111]. This study also found evidence for increased methylation patterns caused by viral infection in one of the barley samples investigated. In this manner, methylation patterns in ancient DNA can function as a proxy for the stress the organism is under, for example stress induced by climate change. Comparisons of DNA methylation patterns in Denisovans, Neanderthals and modern humans have revealed differentially methylated regions in genes relating to disease and limb development, providing clues into the phenotypic differences between archaic and modern humans [112]. Provided that genome coverage is sufficiently high, epigenomic information can be bioinformatically deciphered directly from the ancient reads (provided appropriate adaptation of the lab-protocols are made to retain the methylation patterns), without the need for additional DNA extraction, or sample destruction, and thus may become part of the standard 'tool-kit' of ancient genomics.

Conclusions and outlook

Technological and intellectual developments have greatly increased the accessibility of aDNA, and have expanded the potential for novel research. Many of the available techniques are already being adapted to suit the specific needs of aDNA, but efforts in the lab should continue to focus on comparing and optimising protocols for all steps of the analysis. Particularly in the field of hybridisation capture, there is much room for technological development. Many applications also assume some level of available sequence information, such as mapping of NGS reads to a reference, and are therefore expected to become more accurate as more sequence information is gathered for extinct populations and species. Finally, although categorical thinking may be convenient for the human mind, if aDNA researchers want to make full use of the insights that palaeogenomic data can offer, they need to adopt a population genetics view that allows the incorporation of all the blurry gradations of real world processes. For example, the traditional phylogenetic tree may have to be replaced by more complex but also more realistic representations of the evolutionary process incorporating genetic admixture and isolation (e.g. TreeMix) that reflect the true population history more accurately.

Acknowledgements

A. L. is supported by DFG grant No. DFG 852/7-4, C. F. S. by Marie-Curie grant ORCA FP7-PEOPLE-2011-IIF 299075 and M. H. by ERC consolidator grant No. 310763 GeneFlow.

References

1. **Pääbo S**. 1985. Preservation of DNA in ancient Egyptian mummies. *J Archaeol Sci* **12**: 411–7.
2. **Higuchi R, Bowman B, Freiberger M, Ryder OA**, et al. 1984. DNA sequences from the quagga, an extinct member of the horse family. *Nature* **312**: 282–4.
3. **Pääbo S**. 1985. Molecular cloning of ancient Egyptian mummy DNA. *Nature* **314**: 644–5.
4. **Leonard JA, Rohland N, Glaberman S, Fleischer RC**, et al. 2005. A rapid loss of stripes: the evolutionary history of the extinct quagga. *Biol Lett* **1**: 291–5.
5. **Pääbo S, Wilson AC**. 1988. Polymerase chain reaction reveals cloning artefacts. *Nature* **334**: 387–8.
6. **Pääbo S, Poinar H, Serre D, Jaenicke-Després V**, et al. 2004. Genetic analyses from ancient DNA. *Annu Rev Genet* **38**: 645–79.
7. **Gilbert MTP, Bandelt H-J, Hofreiter M, Barnes I**. 2005. Assessing ancient DNA studies. *Trends Ecol Evol* **20**: 541–4.
8. **Lambert DM, Ritchie PA, Millar CD, Holland B**, et al. 2002. Rates of evolution in ancient DNA from Adélie penguins. *Science* **295**: 2270–3.
9. **Hofreiter M, Serre D, Poinar HN, Kuch M**, et al. 2001. Ancient DNA. *Nat Rev Genet* **2**: 353–9.
10. **Paijmans JLA, Gilbert MTP, Hofreiter M**. 2013. Mitogenomic analyses from ancient DNA. *Mol Phylogenet Evol* **69**: 404–16.
11. **Huynen L, Millar CD, Lambert DM**. 2012. Resurrecting ancient animal genomes: the extinct moa and more. *BioEssays* **34**: 661–9.
12. **Gilbert MTP, Cuccui J, White W, Lynnerup N**, et al. 2004. Absence of *Yersinia pestis*-specific DNA in human teeth from five European excavations of putative plague victims. *Microbiology* **150**: 341–54.
13. **Shapiro B, Rambaut A, Gilbert MTP**. 2006. No proof that typhoid caused the Plague of Athens (a reply to Papagrigorakis et al.). *Int J Infect Dis* **10**: 334–5.
14. **Margulies M, Egholm M, Altman WE, Attiya S**, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376–80.
15. **Mamanova L, Coffey AJ, Scott CE, Kozarewa I**, et al. 2010. Target-enrichment strategies for next-generation sequencing. *Nat Meth* **7**: 111–8.
16. **Briggs AW, Good JM, Green RE, Krause J**, et al. 2009. Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science* **325**: 318–21.
17. **Maricic T, Whitten M, Pääbo S**. 2010. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* **5**: e14004.
18. **Carpenter ML, Buenrostro JD, Valdiosera C, Schroeder H**, et al. 2013. Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. *Am J Hum Genet* **93**: 852–64.
19. **Enk JM, Devault AM, Kuch M, Murgha YE**, et al. 2014. Ancient whole genome enrichment using baits built from modern DNA. *Mol Biol Evol* **31**: 1292–4.
20. **Dabney J, Knapp M, Glocke I, Gansauge M-T**, et al. 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci USA* **110**: 15758–63.
21. **Brotherton P, Endicott P, Sanchez JJ, Beaumont M**, et al. 2007. Novel high-resolution characterization of ancient DNA reveals C > U-type base modification events as the sole cause of post mortem miscoding lesions. *Nucleic Acids Res* **35**: 5717–28.
22. **Briggs AW, Stenzel U, Meyer M, Krause J**, et al. 2010. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Res* **38**: e87.
23. **Briggs AW, Stenzel U, Johnson PLF, Green RE**, et al. 2007. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci USA* **104**: 14616–21.
24. **Orlando L, Ginolhac A, Raghavan M, Vilstrup J**, et al. 2011. True single-molecule DNA sequencing of a Pleistocene horse bone. *Genome Res* **21**: 1705–19.

25. **Sawyer S, Krause J, Guschanski K, Savolainen V**, et al. 2012. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* **7**: e34131.
26. **Ginolhac A, Rasmussen M, Gilbert MTP, Willerslev E**, et al. 2011. mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* **27**: 2153–5.
27. **Jónsson H, Ginolhac A, Schubert M, Johnson PLF**, et al. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**: 1682–4.
28. **Smith O, Clapham A, Rose P, Liu Y**, et al. 2014. A complete ancient RNA genome: identification, reconstruction and evolutionary history of archaeological Barley Stripe Mosaic Virus. *Sci Rep* **4**: 4003.
29. **Gansauge M-T, Meyer M**. 2014. Selective enrichment of damaged DNA molecules for ancient genome sequencing. *Genome Res* **24**: 1543–9.
30. **Skoglund P, Northoff BH, Shunkov MV, Derevianko AP**, et al. 2014. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc Natl Acad Sci USA* **111**: 2229–34.
31. **Poinar HN, Höss M, Bada JL, Pääbo S**. 1996. Amino acid racemization and the preservation of ancient DNA. *Science* **272**: 864–6.
32. **Collins MJ, Penkman KEH, Rohland N, Shapiro B**, et al. 2009. Is amino acid racemization a useful tool for screening for ancient DNA in bone? *Proc R Soc B Biol Sci* **276**: 2971–7.
33. **Allentoft ME, Collins M, Harker D, Haile J**, et al. 2012. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc R Soc B Biol Sci* **279**: 4724–33.
34. **Hughey JR, Braga JC, Aguirre J, Woelkerling WJ**, et al. 2008. Analysis of ancient DNA from fossil corallines (Corallinales, Rhodophyta). *J Phycol* **44**: 374–83.
35. **Hijmans RJ, Cameron S, Parra J, Jones PG**, et al. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* **25**: 1965–78.
36. **Bintanja R, van de Wal RSW, Oerlemans J**. 2005. Modelled atmospheric temperatures and global sea levels over the past million years. *Nature* **437**: 125–8.
37. **Rasmussen M, Anzick SL, Waters MR, Skoglund P**, et al. 2014. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* **506**: 225–9.
38. **Rohland N, Hofreiter M**. 2007. Comparison and optimization of ancient DNA extraction. *BioTechniques* **42**: 343–52.
39. **Bouwman AS, Brown TA**. 2002. Comparison between silica-based methods for the extraction of DNA from human bones from 18th to Mid-19th century London. *Anc Biomol* **4**: 173–8.
40. **Yang DY, Eng B, Waye JS, Dudar JC**, et al. 1998. Improved DNA extraction from ancient bones using silica-based spin columns. *Am J Phys Anthropol* **105**: 539–43.
41. **Meyer M, Fu Q, Aximu-Petri A, Glocke I**, et al. 2013. A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* **505**: 403–6.
42. **Orlando L**. 2014. A 400,000-year-old mitochondrial genome questions phylogenetic relationships amongst archaic hominins. *BioEssays* **36**: 598–605.
43. **Orlando L, Ginolhac A, Zhang G, Froese D**, et al. 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**: 74–8.
44. **Geigl E-M**. 2002. On the circumstances surrounding the preservation and analysis of very old DNA. *Archaeometry* **44**: 337–42.
45. **Barta JL, Monroe C, Teisberg JE, Winters M**, et al. 2014. One of the key characteristics of ancient DNA, low copy number, may be a product of its extraction. *J Archaeol Sci* **46**: 281–9.
46. **Gilbert MTP, Tomsho LP, Rendulic S, Packard M**, et al. 2007. Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. *Science* **317**: 1927–30.
47. **Rasmussen M, Li Y, Lindgreen S, Pedersen JS**, et al. 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463**: 757–62.
48. **Gamba C, Jones ER, Teasdale MD, McLaughlin RL**, et al. 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nat Commun* **4**: 5257.
49. **Noonan JP, Hofreiter M, Smith D, Priest JR**, et al. 2005. Genomic sequencing of Pleistocene cave bears. *Science* **309**: 597–9.
50. **Green RE, Krause J, Ptak SE, Briggs AW**, et al. 2006. Analysis of one million base pairs of Neanderthal DNA. *Nature* **444**: 330–6.
51. **Briggs AW, Heyn P**. 2012. Preparation of next-generation sequencing libraries from damaged DNA. In Shapiro B, Hofreiter M, eds; *Ancient DNA*. New York: Humana Press. pp. 143–54.
52. **Maricic T, Pääbo S**. 2009. Optimization of 454 sequencing library preparation from small amounts of DNA permits sequence determination of both DNA strands. *BioTechniques* **46**: 51–2, 54–7.
53. **Bennett EA, Massilani D, Lizzo G, Daligault J**, et al. 2014. Library construction for ancient genomics: single strand or double strand? *BioTechniques* **56**: 289–90, 292–6, 298, passim.
54. **Meyer M, Kircher M, Gansauge M-T, Li H**, et al. 2012. A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**: 222–6.
55. **Gansauge M-T, Meyer M**. 2013. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat Protoc* **8**: 737–48.
56. **Branton D, Deamer DW, Marziali A, Bayley H**, et al. 2008. The potential and challenges of nanopore sequencing. *Nat Biotechnol* **26**: 1146–53.
57. **Eid J, Fehr A, Gray J, Luong K**, et al. 2009. Real-time DNA sequencing from single polymerase molecules. *Science* **323**: 133–8.
58. **Skoglund P, Malmström H, Raghavan M, Storå J**, et al. 2012. Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336**: 466–9.
59. **Zhang H, Paijmans JLA, Chang F, Wu X**, et al. 2013. Morphological and genetic evidence for early Holocene cattle management in northeastern China. *Nat Commun* **4**: 2755.
60. **Burbano HA, Hodges E, Green RE, Briggs AW**, et al. 2010. Targeted investigation of the Neandertal genome by array-based sequence capture. *Science* **328**: 723–35.
61. **Devault AM, Golding GB, Waglechner N, Enk JM**, et al. 2014. Second-pandemic strain of *Vibrio cholerae* from the Philadelphia cholera outbreak of 1849. *N Engl J Med* **370**: 334–40.
62. **Enk J, Rouillard J-M, Poinar H**. 2013. Quantitative PCR as a predictor of aligned ancient DNA read counts following targeted enrichment. *BioTechniques* **55**: 300–9.
63. **Castellano S, Parra G, Sánchez-Quinto FA, Racimo F**, et al. 2014. Patterns of coding variation in the complete exomes of three Neandertals. *Proc Natl Acad Sci USA* **111**: 6666–71.
64. **Cosart T, Beja-Pereira A, Chen S, Ng SB**, et al. 2011. Exome-wide DNA capture and next generation sequencing in domestic and wild species. *BMC Genomics* **12**: 347.
65. **Bi K, Vanderpool D, Singhal S, Linderoth T**, et al. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics* **13**: 403.
66. **Fortes GG, Speller CF, Hofreiter M, King TE**. 2013. Phenotypes from ancient DNA: approaches, insights and prospects. *BioEssays* **35**: 690–5.
67. **Elhaik E, Greenspan E, Staats S, Krahn T**, et al. 2013. The GenoChip: a new tool for genetic anthropology. *Genome Biol Evol* **5**: 1021–31.
68. **Charlier C, Coppieters W, Rollin F, Desmecht D**, et al. 2008. Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nat Genet* **40**: 449–54.
69. **Martinsohn JT, Ogden R**. 2009. FishPopTrace—Developing SNP-based population genetic assignment methods to investigate illegal fishing. *Forensic Sci Int Genet Suppl Ser* **2**: 294–6.
70. **Backström N, Fagerberg S, Ellegren H**. 2008. Genomics of natural bird populations: a gene-based set of reference markers evenly spread across the avian genome. *Mol Ecol* **17**: 964–80.
71. **Decker JE, Pires JC, Conant GC, McKay SD**, et al. 2009. Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. *Proc Natl Acad Sci USA* **106**: 18644–9.
72. **Nielsen R**. 2004. Population genetic analysis of ascertained SNP data. *Hum Genomics* **1**: 218.
73. **Clark AG, Hubisz MJ, Bustamante CD, Williamson SH**, et al. 2005. Ascertainment bias in studies of human genome-wide polymorphism. *Genome Res* **15**: 1496–502.
74. **Schubert M, Ginolhac A, Lindgreen S, Thompson JF**, et al. 2012. Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* **13**: 178.
75. **Green RE, Malaspinas A-S, Krause J, Briggs AW**, et al. 2008. A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* **134**: 416–26.
76. **Hahn C, Bachmann L, Chevreux B**. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res* **41**: e129.
77. **Rajaraman A, Tannier E, Chauve C**. 2013. FPSAC: fast phylogenetic scaffolding of ancient contigs. *Bioinformatics* **29**: 2987–94.

78. **Haile J, Holdaway R, Oliver K, Bunce M**, et al. 2007. Ancient DNA chronology within sediment deposits: are paleobiological reconstructions possible and is DNA leaching a factor? *Mol Biol Evol* **24**: 982–9.
79. **Wood DE, Salzberg SL**. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* **15**: R46.
80. **Skoglund P, Sjödin P, Skoglund T, Lascoux M**, et al. 2014. Investigating population history using temporal genetic differentiation. *Mol Biol Evol* **31**: 2516–27.
81. **Pickrell JK, Pritchard JK**. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet* **8**: e1002967.
82. **Sites JW, Marshall JC**. 2004. Operational criteria for delimiting species. *Annu Rev Ecol Evol Syst* **35**: 199–227.
83. **Queiroz KD**. 2007. Species concepts and species delimitation. *Syst Biol* **56**: 879–86.
84. **Reich D, Patterson N, Kircher M, Delfin F**, et al. 2011. Denisova admixture and the first modern human dispersals into southeast Asia and Oceania. *Am J Hum Genet* **89**: 516–28.
85. **Hellenthal G, Busby GBJ, Band G, Wilson JF**, et al. 2014. A genetic atlas of human admixture history. *Science* **343**: 747–51.
86. **Pardini AT, Jones CS, Noble LR, Kreiser B**, et al. 2001. Sex-biased dispersal of great white sharks. *Nature* **412**: 139–40.
87. **Frantz LA, Schraiber JG, Madsen O, Megens H-J**, et al. 2013. Genome sequencing reveals fine scale diversification and reticulation history during speciation in *Sus*. *Genome Biol* **14**: R107.
88. **Rohde DLT, Olson S, Chang JT**. 2004. Modelling the recent common ancestry of all living humans. *Nature* **431**: 562–6.
89. **Pickrell J, Reich D**. 2014. Towards a new history and geography of human genes informed by ancient DNA. *Trends Genet* **30**: 377–89.
90. **Fordyce SL, Ávila-Arcos MC, Rasmussen M, Cappellini E**, et al. 2013. Deep sequencing of RNA from ancient maize kernels. *PLoS One* **8**: e50961.
91. **Sallon S, Solowey E, Cohen Y, Korchinsky R**, et al. 2008. Germination, genetics, and growth of an ancient date seed. *Science* **320**: 1464.
92. **Santiago-Rodriguez TM, Patrício AR, Rivera JI, Coradin M**, et al. 2014. luxS in bacteria isolated from 25- to 40-million-year-old amber. *FEMS Microbiol Lett* **350**: 117–24.
93. **Weyrich LS, Llamas B, Cooper A**. 2014. Reply to Santiago-Rodriguez et al.: was luxS really isolated from 25- to 40-million-year-old bacteria? *FEMS Microbiol Lett* **353**: 85–6.
94. **Warinner C, Rodrigues JFM, Vyas R, Trachsel C**, et al. 2014. Pathogens and host immunity in the ancient human oral cavity. *Nat Genet* **46**: 336–44.
95. **Achtman M**. 2012. Insights from genomic comparisons of genetically monomorphic bacterial pathogens. *Philos Trans R Soc B Biol Sci* **367**: 860–7.
96. **Campana MG, García NR, Rühli FJ, Tuross N**. 2014. False positives complicate ancient pathogen identifications using high-throughput shotgun sequencing. *BMC Res Notes* **7**: 111.
97. **Warinner C, Speller C, Collins M**. 2015. A new era in paleomicrobiology: Prospects for ancient dental calculus as a long-term record of the human oral microbiome. *Philos Trans R Soc B Biol Sci*, in press, doi: 10.1098/rstb.2013.0376.
98. **Adler CJ, Dobney K, Weyrich LS, Kaidonis J**, et al. 2013. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nat Genet* **45**: 450–5.
99. **Tito RY, Knights D, Metcalf J, Obregon-Tito AJ**, et al. 2012. Insights from characterizing extinct human gut microbiomes. *PLoS One* **7**: e51146.
100. **Maixner F, Thomma A, Cipollini G, Widder S**, et al. 2014. Metagenomic analysis reveals presence of *Treponema denticola* in a tissue biopsy of the iceman. *PLoS One* **9**: e99994.
101. **Chapman SJ, Hill AVS**. 2012. Human genetic susceptibility to infectious disease. *Nat Rev Genet* **13**: 175–88.
102. **Thalmann O, Shapiro B, Cui P, Schuenemann VJ**, et al. 2013. Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science* **342**: 871–4.
103. **Ludwig A, Pruvost M, Reissmann M, Benecke N**, et al. 2009. Coat color variation at the beginning of horse domestication. *Science* **324**: 485.
104. **Flink LG, Allen R, Barnett R, Malmström H**, et al. 2014. Establishing the validity of domestication genes using DNA from ancient chickens. *Proc Natl Acad Sci USA* **111**: 6184–9.
105. **Larson G, Piperno DR, Allaby RG, Purugganan MD**, et al. 2014. Current perspectives and the future of domestication studies. *Proc Natl Acad Sci USA* **111**: 6139–46.
106. **Tishkoff SA, Reed FA, Ranciaro A, Voight BF**, et al. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat Genet* **39**: 31–40.
107. **Itan Y, Powell A, Beaumont MA, Burger J**, et al. 2009. The Origins of Lactase Persistence in Europe. *PLoS Comput Biol* **5**: e1000491.
108. **Avise JC**. 2000. *Phylogeography: The History and Formation of Species*. Cambridge MA: Harvard University Press.
109. **Hauser L, Adcock GJ, Smith PJ, Ramírez JHB**, et al. 2002. Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proc Natl Acad Sci USA* **99**: 11742–7.
110. **Pedersen JS, Valen E, Velazquez AMV, Parker BJ**, et al. 2014. Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. *Genome Res* **24**: 454–66.
111. **Smith O, Clapham AJ, Rose P, Liu Y**, et al. 2014. Genomic methylation patterns in archaeological barley show de-methylation as a time-dependent diagenetic process. *Sci Rep* **4**: 5559.
112. **Gokhman D, Lavi E, Prüfer K, Fraga MF**, et al. 2014. Reconstructing the DNA methylation maps of the Neandertal and the Denisovan. *Science* **344**: 523–7.