What (molecular) time is it?

Using ancient DNA to date evolutionary events

Julien Soubrier

Australian Centre for Ancient DNA
Ecology and Evolutionary Biology
School of Earth and Environmental Sciences
The University of Adelaide

A thesis submitted for the degree of Doctor of Philosophy at the University of Adelaide

April 2012

Table of contents

Abstract

This work aims to explore the use of genetic sequences sampled serially through time (heterochronous data), to infer the timescale of past evolutionary events. Such data can be generated from preserved sub-fossil or fossil organismal remains (like mummified tissues, fossilized bones or coprolites), and then used to observe genetic modifications in real-time. Most importantly, the dates of the samples provide firm temporal tie points for their genetic sequences, and can be used to calibrate phylogenetic reconstructions.

This thesis presents several case studies where ancient DNA was used to re-calibrate evolutionary timescales. In every situation, the use of heterochronous data led to elevated molecular rate estimates, resulting in the reconstruction of younger timescales, as compared to estimates based on fossil calibrations. These observations are in agreement with the recent demonstration that molecular rates vary according to the time period over which they are calculated.

This work shows that, ancient DNA offers crucial temporal information to reliably estimate the timescale of recent population evolution, and is generally the only source of direct calibration available for this specific timeframe.

Along with the results specific to each organism studied (hyena, bison and human), an emphasis was placed on the methodological aspects of the use of ancient DNA to generate timed phylogenetic inferences. Additionally, simulated data and mathematical modelling were used to extend the understanding of specific aspects of the temporal dependence of molecular rates.

The results discussed in the present study help to further elucidate the evolutionary mechanisms behind the molecular clock concept, and have implications for the development and application of statistical models to obtain accurate time estimates from genetic data.

Thesis declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution to Julien Soubrier and, to the best of my knowledge and belief, contains no material previously published or written by any other person, except where due reference has been made in the text.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis (as listed below) resides with the copyright holders(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue, the Australian Digital Theses Program (ADTP) and also through web research engines, unless permission has been granted by the University to restrict access for a period of time.

The published works of my thesis include:

Ho, SYW, R Lanfear, L Bromham, MJ Phillips, J Soubrier, AG Rodrigo, A Cooper. 2011. Time-dependent rates of molecular evolution. Molecular ecology 20:3087 3101.

Soubrier, J, M Steel, MSY Lee, C Der Sarkissian, S Guindon, SYW Ho, A Cooper. 2012. The influence of rate heterogeneity among sites on the time dependence of molecular rates. Molecular Biology and Evolution 10.1093

_____

Julien Soubrier

20th April 2012

Acknowledgements

I would like to thank my principal supervisor Alan Cooper for giving me the opportunity to work at the Australian Centre for Ancient DNA, his constant challenging scientific discussions, and his kind support throughout the three years of my PhD candidature in Australia. I am very grateful to my co-supervisors, Simon Ho, Wolfgang Haak and Mike Lee, for their enthusiastic guidance and encouragements.

I would like to recognize all fellow members at the ACAD for providing such a friendly working environment, where passion for science and a sharing of knowledge go along with after-hour events. In particular, all the thesis writing group fellows were a great support.

I would like to show my gratitude to Vicki Thomson for her unceasing technical, intellectual and friendly support during the long hours we spent at work trying to simultaneously finish our theses (but also trying to get the best coffee in the lab); Bastien Llamas for his endless enthusiasm and inspiration at sorting illustrations, presentations, paragraphs and shell scripts, but mostly for his friendship; and Maria Lekis for her invaluable support throughout my PhD.

This work would not have been possible without the financial and intellectual support of The Genographic Project, in collaboration with the National Geographic, IBM and the Waitt Family Foundation.

I would like to thank all co-authors of the manuscripts presented in this thesis, for their support and helpful comments.

I would also like to acknowledge the people who spent time in discussions, and explanations during conferences, workshops or visits, notably Mike Steel, Alexei Drummond, Stephane Guindon, David Bryant, David Penny, Jerry Taylor, Mark Pagel, Rob Lanfear, Matt Philipps, Lindell Bromham and Chris Simon.

I am extremely grateful to Christophe Douady and Manolo Gouy for teaching me a lot about scientific research during my Masters in Lyon; Sandrine Hughes for training me in ancient DNA techniques; and Ludovic Orlando for introducing me to this fascinating field of paleogenetics, and for his helpful advice on studies and career choices.

I would also like to thank Profs. Pierre Joly and Michel Aigle for the passion they put in their lectures, which were decisive moments for my choosing to continue in research.

Finally, on a more personal note, I would like to thank the people, in and out of the lab, that made me feel welcome in this foreign environment, and helped transform this time in Australia into an amazing experience. Among them,

Bastien, Emmanuelle, Clio, Camille, Damien, Virginie, Rethish, Hugh, Patricia, Sarah, Doreen, Kimi and Jess. I am heartily grateful to Emma, Yves, Chris, Sacha, Bea, Quentin and all the Apoptose group from the master, for being such great friends regardless of the distance, with a special thanks to Adrien. I can't find the words to express the gratitude I feel for my wife for following me in this adventure, for her support and patience, and for making the past two years the most joyful moment of my life by giving birth to an amazing little boy, Noah, who always manages to make me smile at the end of the day. I owe my deepest gratitude to my family for inestimable support, their encouragement and their visits to Australia, my parents Marc and Françoise, but also Evelyne, Pierre, Jean-Pierre, Clara, André, Mimi, Françoise, Marilou, Clément, Paulette, Bernard and Evelyne, with a immense thank you to Laurence.

Chapter 1:

Introduction

Inferring the date of past evolutionary events from genetic sequences has remained a challenge since the introduction of the molecular clock hypothesis by Zuckerkandl and Pauling 50 years ago (Zuckerkandl, Pauling 1962). From this simple concept, that changes in molecular sequences occur at a constant rate through time, models have developed in complexity to better match the biological realities of molecular evolution.

In parallel, considerable technological progress has been made in methods to acquire genetic information, reaching the point where it is possible to sequence complete genomes from virtually any complex organism. The availability of an increasing amount of data is particularly beneficial to the field of evolutionary biology, with a vast increase in the number of characters to compare between taxa and an improvement in taxon sampling. The advent of next generation sequencing methods has also been a major advantage for paleogeneticists, and has led to the recent sequencing of mitochondrial and nuclear genomes of several extinct taxa, like Neandertal (Briggs et al. 2009; Green et al. 2010), mammoths (Gilbert et al. 2008; Miller et al. 2008), bears (Noonan et al. 2005; Lindqvist et al. 2010), or anatomically modern humans (Krause et al. 2010; Rasmussen et al. 2010), (see also Ho, Gilbert 2010; Stoneking, Krause 2011, as reviews).

In the context of molecular dating, these ancient sequences are of particular importance due to the temporal information they contain. Until recently, most molecular clock studies have been based on the fossil record, which requires the use of estimated divergence dates between species or group of species, often in the order of magnitude of several million years. But it is now well established that molecular rates vary according to the time period over which they are calculated, and therefore, fossil calibrations are problematic for inferring recent evolutionary events such as population-level studies. For that reason, the dates associated with molecular sequences that have been sampled serially through time constitute a precious source of calibrations within the timescale of recent evolutionary history.

This thesis aims to further explore the nature and extent of biases related to the temporal dependence of molecular rates, and present case studies of the utility of ancient DNA for the reconstruction of phylogenetic timescales. To achieve this, seven manuscripts are compiled into five chapters:

- The first study (Chapter 2) is a re-assessment of the spotted hyena evolutionary timescale, using ancient DNA sequences.
- The second and third studies are focusing on bison evolution, by using ancient and modern DNA either from mitochondrial markers (Chapter 3), or from genome-wide nuclear markers (Chapter 4).
- The two articles composing Chapter 5, aim at better understanding the issue of temporal dependence of molecular rates: a review of the subject, followed by an

exploration of one of its potential causes using mathematical models and simulated data.

- Finally, two ancient DNA studies of modern humans are compiled in the last chapter (Chapter 6). These studies are based on the first dataset of ancient whole mitochondrial genomes sequenced at a population level. They were performed within the framework of The Genographic Project, a worldwide research initiative to genetically characterize the recent evolution of human populations.

Phylogenetics

Molecular markers

Phylogenetic reconstructions aim to infer the evolutionary history of organisms based on the comparison of homologous characters between individuals.
Sequences of proteins, or DNA from nuclear or mitochondrial genomes, are most commonly used to infer organismal phylogenetic trees. The taxonomic rank between individuals of interest often dictates the choice of molecular marker to use, with slowly evolving protein sequences more commonly used for deep phylogenies, while rapidly evolving mitochondrial control region sequences are often used for population-level studies. But the many properties of different molecular markers available need to be fully considered.

For example, the mitochondrial genome is only maternally inherited and is non-recombinant, allowing evolutionary history to be traced back directly by following mutations between generations (Atkinson, Gray, Drummond 2008; Galtier et al. 2009). In addition, mitochondrial DNA is easily amplifiable due to its presence in multiple copies per cell, and therefore constitutes a target of choice for ancient DNA studies (Ho, Gilbert 2010). On the other hand, some of these properties can also be limitations: the lack of recombination means that the entire mitochondrial genome is effectively acting as a single locus, and phylogenies based on mitochondrial markers do not necessarily represent accurately the evolutionary history of the organisms of interest.

As organismal evolution is the result of genetic changes throughout the entire genome (nuclear and mitochondrial), a mixture of different loci can be used in order to obtain more precise phylogenetic inferences. Currently, it is not computationally feasible to perform phylogenetic analyses of complete genomes where the genomes are large and complex. But because many genomic regions are not phylogenetically informative, the use of single-nucleotide polymorphism (SNP) chip methods offer a solution by genotyping a selection of variable sites widely distributed at a genomic level (LaFramboise 2009). However, SNP chips have mainly been developed for purposes other than evolutionary studies, and therefore their

specificities (including genotype coding or SNP selection) must be taken into account before applying traditional phylogenetic tools. The specificities of SNP chips and their implications for phylogenetic reconstruction are explored in chapter 4, using the phylogenetic relationships of bison as a case study.

Phylogenetic methods

Early methods of phylogenetic reconstruction such as Neighbour joining (using genetic distances), or maximum parsimony (directly using molecular characters), have paved the way for more complex probabilistic methods: maximum likelihood and Bayesian approaches. The statistical framework of such probabilistic methods allows explicit substitution models with a large number of parameters to be included in phylogenetic reconstructions in a reasonable amount of time. For example, the use of Bayesian Monte Carlo Markov Chain (MCMC) methods allows the integration of multiple parameters such as demographic processes, temporal information and topological priors within the same calculation (Drummond, Rambaut 2007). These capacities of maximum likelihood and Bayesian methods are used and explored in the different chapters of this thesis (i.e., phylogenetic reconstructions from three-character state data in the study of bison nuclear SNP libraries, Chapter 3).

DNA sequence substitution models

Increasingly complex models have been developed to account for the biological processes behind the different types of nucleotide character changes. For example, Jukes-Cantor (Jukes, Cantor 1969) is the simplest substitution model, and assumes that all nucleotides are at the same frequency in the data and that all possible types of substitution have equal probabilities. With more complex models, different degrees of heterogeneity are allowed, either between substitution types, or between base frequencies. The General Time Reversible model (Rodriguez et al. 1990) allows for both types of heterogeneity.

It is important to note that, due to the complexity of the biological rules underlying the mutation processes, no mathematical model can perfectly describe molecular sequence evolution, and consequently none of the models are likely to be biologically correct (some signal is lost when the evolution of the different genetic characters is mathematically modelled). In theory, the more complex the model, the better it is likely to fit real data. But with a limited number of characters available, the more parameters are considered in the model, the less data there is per parameter (Posada, Buckley 2004). Based on that principle, statistical tests have been developed to help select the most appropriate substitution model for a particular dataset, balancing the fitness of the data against the complexity of the model. The

most commonly used tests are the Akaike Information Criterion (AIC) 1 and 2, and the Bayesian Information Criterion (Posada, Buckley 2004; Keane et al. 2006). The AIC tests are based on likelihood statistics, and allow the quantification of the amount of genetic signal lost when using a specific model of evolution, which is a function of the number of parameters estimated by the model. The model with the least information lost is therefore selected. The BIC is based on the comparison Bayesian posterior probabilities. The most noticeable difference between these 3 tests (AIC1, AIC2 and BIC) is their degree of handicap against the addition of extra parameters. In that regard, AIC2 and BIC penalize more complex models, and therefore limit over-parameterization.

Data partitioning

For short molecular sequences (like the mitochondrial control region), the tests described above provide a simple solution to empirically select an appropriate substitution model. But for longer sequences (from a few genes to genome-wide data sets) a single model cannot take into account the high level of heterogeneity in mutation patterns generated by natural processes, such as differential selection pressure between genomic regions. To reduce the amount of heterogeneity among characters, the sequence data can be divided into subsets (i.e., partitioned), and different substitution models can be estimated independently for each subset. In some cases, the patterns of selection pressure is well described (for example between 1^{st}, 2^{nd}, and 3^{rd} codon positions in protein coding sequences) and therefore define obvious partitioning schemes. But other factors such as base composition, or rate heterogeneity among sites, should also be taken into account when partitioning the data, leading to a large number of potential partitioning schemes. Similarly to the situation above concerning model selection, there is a balance between increasing the number of subsets to better describe the data, and over-parameterization where the amount of signal present in each subset is reduced. Based on this observation, Lanfear and co-authors recently developed a method to extract the best partitioning scheme empirically from a sequence alignment (Lanfear et al. 2012). This method is combined with the model selection tests mentioned above in a program called PartitionFinder, which was used in Chapter 6 to partition human whole mitochondrial genome datasets.

The molecular clock

Assuming a constant mutation rate through time (a strict molecular clock), it is possible to date past events, like speciation events or the time of the most recent common ancestor (tMRCA) of a clade, by comparing the present genetic diversity to a calibration point. For example, if 10 genetic differences are observed within a 100 base pair (bp)

sequence between two taxa known to have been separated for 1 million years (My), a molecular rate of 0.1 substitutions per site per million years (s/s/My) is calculated. This rate can then be used to examine a situation where the same sequence presents 2 differences between another pair of taxa, leading to an estimation of 200,000 years (ky) since their divergence.

The first time genetic mutations were assumed to occur at a relatively constant rate was within protein sequences by Zuckerkandl and Pauling in 1962, in order to estimate the divergence time of different globin types (Zuckerkandl, Pauling 1962). Later, Kimura (1968; 1983) theorized the concept of strict molecular clocks based on the assumption that a vast majority of genetic changes are neutral: most mutations do not drastically affect the fitness of an organism, therefore their fixation or disappearance is mainly due to genetic drift and not natural selection. This 'neutral theory' consequently implicates that, regardless of population size or generation time, the long-term substitution rate observed between organisms is similar to the mutation rate of the DNA molecule between generations.

Due to the development of DNA nucleotide amplification and sequencing methods, the molecular clock concept has rapidly been challenged, and large discrepancies have been reported between molecular rate estimates (e.g. Wu, Li 1985; Britten 1986; Gaut et al. 1992)

Rate heterogeneity among sites and among lineages

Two principal sources of variance from a strict molecular clock were first used to explain the differences observed between time estimates: (1) rate heterogeneity among sites and (2) rate heterogeneity among lineages (Arbogast et al. 2002; Pulquerio, Nichols 2007).

(1) Among-site rate heterogeneity refers to a situation where not all sites of an alignment are susceptible to change at the same rate, mainly due to the differential selection pressure between sites. To account for such heterogeneity in the data, Yang (1994; 1996) has shown that a gamma distribution could be used to discretise a continuous distribution of site-specific rates into a pre-defined number of rate categories of equal probability, where the mean is set to one and the variance characterizes the rate variability. The probability of each site being associated with a rate category, and the shape parameter of the distribution are both estimated from the data. A discretized gamma distribution allows the characterization of rate heterogeneity among sites to be empirically estimated for each dataset, and integrated into the analysis within a reasonable computational requirement. However, this method has two main limitations: it is based on the assumption that all situations of among-site rate heterogeneity can be modelled with a single gamma-like distribution; and the discretization step reduces the

full range of site specific rates to a small number of rate categories (where all rate values of the category are approximated by the mean rate of the category).

(2) Among-lineage rate heterogeneity describes the fact that mutation rates may differ between taxa. To account for such heterogeneity, algorithms have been developed to relax the clock assumption among the branches of phylogenetic trees (Thorne, Kishino, Painter 1998; Huelsenbeck, Larget, Swofford 2000). Such methodological developments were possible due to the implementation of probabilistic tools (like maximum likelihood or Bayesian methods) for phylogenetic calculations. Relaxed clock inferences were first based on the assumption that rate changes are continuous and auto-correlated among branches (the rate at one branch depends of the rate of the parent branch), because of a putative inheritance of mutation rates between lineages. Later, it was shown that the auto-correlation of evolutionary rates between lineages was not significant for large data sets and an uncorrelated relaxed clock model was consequently developed (Drummond et al. 2006). More recently, a Random Local Clock (RLC) was introduced (Drummond, Suchard 2010), based on the principle that a discrete number of local strict clocks could be defined between clades of a tree to account for among-lineage rate heterogeneity. The latter two methods can also be used to empirically test the null hypothesis of a strict molecular clock for a sequence alignment: a strict clock cannot be rejected if the posterior distribution of the standard deviation of the estimated rate comprises 0 for the uncorrelated relaxed clock, or if the number of estimated local clocks is 1. The software package BEAST (Drummond, Rambaut 2007) was used to conduct the dated inferences presented in this thesis using the three different clock models: strict clock, uncorrelated relaxed clock, and random local clock.

Different tree models for intra and inter-specific data

In order to reconstruct the evolutionary history of populations (i.e., genealogies), the genetic diversity between individuals can be compared using the coalescent theory as a statistical framework (Kingman 1982; Donnelly, Tavare 1995). Assuming a random selection of individuals from a panmictic population (i.e., with random mating) and moving backwards in time generation by generation, individual haplotypes start to coalesce into ancestral lineages, until a single ancestor is reached (the most recent common ancestor). The coalescent theory can also be used to infer past population demography, by estimating the effective population size through time from a genealogy (Drummond et al. 2005). Methods, such as the skyline plot implemented in BEAST, allow the co-estimation of genealogies and demographic histories from genetic sequences (Ho, Shapiro 2011).

However, when performing phylogenetic reconstruction from samples representing different species, each split of the phylogeny correspond to a speciation event (and not to a coalescent event as it is the case for genealogies). In such situation a birth-death model, or birth only (like the Yule model, Yang, Rannala 1997), can be used as a prior on the tree reconstruction.

In this context, the past evolutionary history of organisms can only be modelled when a single locus sequence (e.g., one gene, or a mitochondrial genome) is collected, either from random individuals of one population, or from single representatives from different species. But this is rather limiting. Being able to combine genetic data from different loci, and from both intra and inter-specific diversity, would allow more precise phylogenetic reconstructions. Luckily, a multispecies coalescent method, called *BEAST ('StarBEAST'), was recently developed, implementing models for both within and between-species genetic data, and combining different gene genealogies into one species tree (Heled, Drummond 2010). The *BEAST option was used in Chapter 6, to accommodate sequences from modern human and chimpanzee in the same phylogenetic analysis.

Further clock inconsistencies

In a review on the accuracy of molecular dating, Pulquerio and Nichols (2007) point out severe inconsistencies between time estimates based on the molecular clock. At that time, such inconsistencies could be attributed to the unreliability of the current methods used to infer dated trees, issues with calibration points and the failure to account for uncertainty.

Subsequently, Ho and co-authors (2008) used newly developed Bayesian methods to illustrate the impact of inappropriate calibrations on molecular date estimates. In addition to the importance of including calibration errors in dated analyses, the authors demonstrate the consequences of using calibration points distant from the date(s) of interest. More precisely, using inter-specific calibration points (like speciation dates inferred from the fossil record) to estimate molecular dates at intra-specific levels (such as population events), or the reciprocal situation, leads to unreliable results.

Temporal dependence of molecular rates

The main reason for this observed clock bias is the apparent time dependence of molecular rates. Rates estimated over short periods of time (e.g., pedigree or genealogy studies) are systematically higher than rates estimated over long periods of time (e.g., species phylogeny) (Ho et al. 2005). A number of factors, such as selection, calibration errors or model misspecifications, have been suggested to explain this apparent temporal dependency.

The first part of Chapter 5 is a detailed review of these potential causes (published in Molecular Ecology as Ho et al. 2011a).

From our review of existing literature on the subject, it became apparent that the effect of saturation (multiple changes occurring at the same site), and rate heterogeneity among sites, on the time-rate estimates was yet to be explored. For that reason, the article presented in chapter 5 aims to show the time-dependent rate pattern resulting from inaccurate modelling of among-site rate heterogeneity (such as the presence of fast-evolving 'hotspots' in an alignment), using mathematical models and simulated data sets.

Ideally, such time-dependent patterns could be characterized empirically and thus corrected. But for molecular rate estimates, a precise characterization of the pattern would involve the availability of accurate calibration points through time as a starting point.

The role of ancient DNA in calibrating dated phylogenies

The use of heterochronous sequences (sampled serially through time) offers one solution to the issue of calibration. When ancient DNA can be amplified and sequenced, the dates associated with each sample (e.g., radiocarbon dates) can be used as primary calibration points (Drummond et al. 2002; Lambert et al. 2002; Shapiro et al. 2004). Such calibration points are of particular importance in the context of the temporal dependence of molecular rates, because they cover a period of time for which very little temporal information is available. In fact, the identification of calibration points from the fossil record is generally only possible between evolutionarily distinct species, as bone remains need to be clearly identifiable between specimens. For example, a paleontological calibration younger than the human/chimpanzee split (5 to 6 My) is lacking for the calibration of human molecular evolution. On the other extreme, direct observation of genetic changes through time (i.e., pedigree studies) are only possible for a limited number of human generations, and depend on genealogical, sampling and genetic amplification and sequencing techniques. This issue is particularly limiting for complex organisms and more specifically large mammals with long generation times; this is explored in three such mammal species in this thesis: hyenas, bovids and humans (chapters 2, 3/4 and 6, respectively). A major gap in calibration points lies between the paleontological record, with dates in the order of magnitude of millions of years, and the pedigree studies, at a scale of decades. This gap encompasses most of the Pleistocene (2.5 million years ago to 12,000 years ago) and much of the Holocene (12,000 years ago to present day), yet most population genetic studies and the phylogenies of recent species will have crucial dates falling within this time range.

In the past few decades, ancient DNA (aDNA) studies have shown the utility of amplifying genetic markers from 100-100,000 years-old samples to identify extinct species,

ancient population patterns or demographic history (e.g., Barnes et al. 2002; Orlando et al. 2002; Shapiro et al. 2004; Ramakrishnan, Hadly 2009). In addition to the ability to directly observe past (and potentially lost) genetic diversity, aDNA has also been successfully used for calibration in a few recent studies, revealing particularly high substitution rate estimates (Ho, Kolokotronis, Allaby 2007; Hay et al. 2008; Subramanian et al. 2009). These observations confirm the described temporal dependence of molecular rates.

Although the special issues related to most aDNA data sets (i.e., short sequences, high levels of damage) have raised questions about their reliability for estimating molecular rates (Debruyne, Poinar 2009), it has now been established that ancient sequences can contain enough genetic and temporal information to obtain accurate rate and date estimates (Ho et al. 2011b). Most notably, parameters like the relative number of ancient sequences in a data set, their distribution in time and within the topology, have been shown to impact the power of dates associated with ancient samples to calibrate a particular phylogeny. A 'date randomization test', which tests the presence of sufficient signal to calibrate phylogenetic analyses, can be used to check if ancient sequences contain enough information to provide calibration (de Bruyn et al. 2009; Firth et al. 2010; Ho et al. 2011b). This test was used in all studies presented in this thesis when data sets contained ancient sequences.

Other technical limitations exist regarding the use of heterochronous sequences (i.e., sampled serially through time). For example, the accurate dating of fossil remains is only possible up to ~60 ky using accelerated mass spectrometry (AMS) radiocarbon methods, but this time range already provides the possibility of dating important population genetic studies. However, biases of population history inferences have been shown when using traditional population genetics statistics that do not incorporate explicit model for heterochronous data (Depaulis, Orlando, Hanni 2009). Finally, Navascués and Emerson (2009) have pointed out potential biases in rate estimates from Bayesian calculations using ancient sequences, when the data strongly violates prior assumptions of the model. However, the authors emphasize the importance of confidence intervals around molecular rate calculations that include the true rate when the model violation is not extreme. Furthermore the sensitivity of Bayesian methods to non-homogenous demographic history (like strong genetic bottlenecks) is also present in modern data sets, and in some cases, the use of ancient sequences offers the only possibility to directly observe the genetic diversity of a population prior to a demographic event (e.g., Shapiro et al. 2004). With the recent technological capacity to recover genome wide molecular information from ancient samples (by direct sequencing or using SNP genotyping), ancient samples are becoming particularly powerful to study the demographic history of non-homogenous population (e.g., after population bottlenecks) (Mourier et al. 2012).

The work presented herein, aims to use simulated data and empirical studies to extend the understanding of specific aspects of the temporal dependence of molecular rates, and the implications of using heterochronous sequences in the reconstruction of evolutionary timescales.

Overview of the project

In the first data chapter (Chapter 2), the timescale of spotted hyena evolution has been re-assessed using ancient sequences as direct internal calibration points. As expected from the apparent temporal dependence of molecular rates, the timescale based on tip calibrations is considerably younger than previous calculations based on fossil calibrations. Encouragingly, the newly calculated timescale reconciles the hyena molecular clock with the paleontological evidence available for the taxon. This chapter shows the strong potential of ancient DNA to provide temporal signals for intraspecific phylogenetic studies, even with short sequences.

In Chapter 3, a survey of ancient bison genetic diversity in Europe is presented. Similar to the hyena study, only short mitochondrial fragments are used, but the bison study differed in sampling depth, both in temporal and geographical distributions. The extensive sampling allowed the characterization of a previously un-described bison clade, and the characterization of a climate-dependent pattern in the evolution of European bison population in the late Pleistocene.

In Chapter 4, the capacity to use commercially-developed nuclear SNP chips on ancient samples is explored, and more generally the methodological implications of conducting phylogenetic inferences from SNP data, using the American bison as a case study.

Chapter 5 is divided in two manuscripts. The first manuscript is a published review article introducing the concept of time-dependent rates of molecular evolution, and reviewing its potential causes. The preparation of this review led to the realization that the responsibility of one particular factor, rate heterogeneity among sites, was not yet fully explored. Consequently, the second part of Chapter 5 (accepted for publication at Molecular Biology and Evolution) combines mathematical and empirical studies of the influence of among-site rate heterogeneity on the time dependence of molecular rates. Among other results, this study shows the importance of data partitioning to accurately infer molecular timescales. Such observations led to the use of the new program, PartitionFinder, to empirically identify a partitioning scheme for the human whole mitochondrial genomes studied in the next chapter.

The final data chapter (Chapter 6) is composed of two human genetics studies using ancient whole mitochondrial genomes. As part of The Genographic Project, 43 complete mitochondrial genomes were sequenced at the Australian Centre for Ancient DNA from

ancient remains of European individuals. Thirty-seven of these genomes (belonging to haplogroup H, the most frequent haplogroup among modern Europeans) are used in the first part of the chapter to refine our understanding of the evolutionary history of modern Europeans. In addition to the characterization of past genetic shifts in ancestral European populations, the study of these ancient sequences revealed the potential of ancient mitochondrial genomes to calibrate the human tree. Consequently, the second part of the chapter presents the first attempt to re-assess the timescale of recent human evolution using the dates associated with ancient sequences as internal calibration points.

References

Arbogast, BS, SV Edwards, J Wakeley, P Beerli, JB Slowinski. 2002. Estimatinf divergence times from molecular data on phylogenetic and population genetic timescales. Annual Review of Ecology and Systematics 33:707-740.

Atkinson, QD, RD Gray, AJ Drummond. 2008. mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. Molecular Biology and Evolution 25:468-474.

Barnes, I, P Matheus, B Shapiro, D Jensen, A Cooper. 2002. Dynamics of Pleistocene population extinctions in Beringian brown bears. Science 295:2267-2270.

Briggs, AW, JM Good, RE Green, et al. 2009. Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science 325:318-321.

Britten, RJ. 1986. Rates of DNA sequence evolution differ between taxonomic groups. Science 231:1393-1398.

de Bruyn, M, BL Hall, LF Chauke, C Baroni, PL Koch, AR Hoelzel. 2009. Rapid response of a marine mammal species to holocene climate and habitat change. PLoS genetics 5:e1000554.

Debruyne, R, HN Poinar. 2009. Time dependency of molecular rates in ancient DNA data sets, a sampling artifact? Systematic biology 58:348-360.

Depaulis, F, L Orlando, C Hanni. 2009. Using classical population genetics tools with heterochroneous data: time matters! Plos One 4:e5541.

Donnelly, P, S Tavare. 1995. Coalescents and genealogical structure under neutrality. Annual review of genetics 29:401-421.

Drummond, AJ, SYW Ho, MJ Phillips, A Rambaut. 2006. Relaxed phylogenetics and dating with confidence. Plos Biology 4:699-710.

Drummond, AJ, GK Nicholls, AG Rodrigo, W Solomon. 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161:1307-1320.

Drummond, AJ, A Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC evolutionary biology 7:214.

Drummond, AJ, A Rambaut, B Shapiro, OG Pybus. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular Biology and Evolution 22:1185-1192.

Drummond, AJ, MA Suchard. 2010. Bayesian random local clocks, or one rate to rule them all. BMC biology 8:114.

Firth, C, A Kitchen, B Shapiro, MA Suchard, EC Holmes, A Rambaut. 2010. Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. Molecular Biology and Evolution 27:2038-2051.

Galtier, N, B Nabholz, S Glemin, GD Hurst. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Molecular ecology 18:4541-4550.

Gaut, BS, SV Muse, WD Clark, MT Clegg. 1992. Relative rates of nucleotide substitution at the rbcL locus of monocotyledonous plants. Journal of Molecular Evolution 35:292-303.

Gilbert, MT, DI Drautz, AM Lesk, et al. 2008. Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes. Proceedings of the National Academy of Sciences of the United States of America 105:8327-8332.

Green, RE, J Krause, AW Briggs, et al. 2010. A draft sequence of the Neandertal genome. Science 328:710-722.

Hay, JM, S Subramanian, CD Millar, E Mohandesan, DM Lambert. 2008. Rapid molecular evolution in a living fossil. Trends in genetics : TIG 24:106-109.

Heled, J, AJ Drummond. 2010. Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution 27:570-580.

Ho, SYW, MT Gilbert. 2010. Ancient mitogenomics. Mitochondrion 10:1-11.

Ho, SYW, SO Kolokotronis, RG Allaby. 2007. Elevated substitution rates estimated from ancient DNA sequences. Biology Letters 3:702-705.

Ho, SYW, R Lanfear, L Bromham, MJ Phillips, J Soubrier, AG Rodrigo, A Cooper. 2011a. Time-dependent rates of molecular evolution. Molecular ecology 20:3087-3101.

Ho, SYW, R Lanfear, MJ Phillips, I Barnes, JA Thomas, SO Kolokotronis, B Shapiro. 2011b. Bayesian estimation of substitution rates from ancient DNA sequences with low information content. Systematic biology 60:366-375.

Ho, SYW, MJ Phillips, A Cooper, AJ Drummond. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Molecular Biology and Evolution 22:1561-1568.

Ho, SYW, U Saarma, R Barnett, J Haile, B Shapiro. 2008. The effect of inappropriate calibration: three case studies in molecular ecology. Plos One 3:e1615.

Ho, SYW, B Shapiro. 2011. Skyline-plot methods for estimating demographic history from nucleotide sequences. Molecular ecology resources 11:423-434.

Huelsenbeck, JP, B Larget, D Swofford. 2000. A compound poisson process for relaxing the molecular clock. Genetics 154:1879-1892.

Jukes, TH, CR Cantor. 1969. Evolution of protein molecules: New York: Academic Press.

Keane, TM, CJ Creevey, MM Pentony, TJ Naughton, JO McLnerney. 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. BMC evolutionary biology 6:29.

Kimura, M. 1968. Evolutionary rate at the molecular level. Nature 217:624-626.

Kimura, M. 1983. The Neutral Theory of Molecular Evolution: Cambridge University Press.

Kingman, JFC. 1982. The coalescent. Stochastic Processes and their Applications 13:235-248.

Krause, J, AW Briggs, M Kircher, T Maricic, N Zwyns, A Derevianko, S Paabo. 2010. A complete mtDNA genome of an early modern human from Kostenki, Russia. Current biology : CB 20:231-236.

LaFramboise, T. 2009. Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. Nucleic acids research 37:4181-4193.

Lambert, DM, PA Ritchie, CD Millar, B Holland, AJ Drummond, C Baroni. 2002. Rates of evolution in ancient DNA from Adelie penguins. Science 295:2270-2273.

Lanfear, R, B Calcott, SYW Ho, S Guindon. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution.

Lindqvist, C, SC Schuster, Y Sun, et al. 2010. Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of polar bear. Proceedings of the National Academy of Sciences of the United States of America 107:5053-5057.

Miller, W, DI Drautz, A Ratan, et al. 2008. Sequencing the nuclear genome of the extinct woolly mammoth. Nature 456:387-390.

Mourier, T, SYW Ho, MTP Gilbert, E Willerslev, L Orlando. 2012. Statistical guidelines for detecting past population shifts using ancient DNA. Molecular Biology and Evolution.

Navascues, M, BC Emerson. 2009. Elevated substitution rate estimates from ancient DNA: model violation and bias of Bayesian methods. Molecular ecology 18:4390-4397.

Noonan, JP, M Hofreiter, D Smith, JR Priest, N Rohland, G Rabeder, J Krause, JC Detter, S Paabo, EM Rubin. 2005. Genomic sequencing of Pleistocene cave bears. Science 309:597-599.

Orlando, L, D Bonjean, H Bocherens, A Thenot, A Argant, M Otte, C Hanni. 2002. Ancient DNA and the population genetics of cave bears (Ursus spelaeus) through space and time. Molecular Biology and Evolution 19:1920-1933.

Posada, D, TR Buckley. 2004. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. Systematic biology 53:793-808.

Pulquerio, MJ, RA Nichols. 2007. Dates from the molecular clock: how wrong can we be? Trends in ecology & evolution 22:180-184.

Ramakrishnan, U, EA Hadly. 2009. Using phylochronology to reveal cryptic population histories: review and synthesis of 29 ancient DNA studies. Molecular ecology 18:1310-1330.

Rasmussen, M, Y Li, S Lindgreen, et al. 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 463:757-762.

Rodriguez, F, JL Oliver, A Marin, JR Medina. 1990. The general stochastic model of nucleotide substitution. Journal of Theoretical Biology 142:485-501.

Shapiro, B, AJ Drummond, A Rambaut, et al. 2004. Rise and fall of the Beringian steppe bison. Science 306:1561-1565.

Stoneking, M, J Krause. 2011. Learning about human population history from ancient and modern genomes. Nature reviews. Genetics 12:603-614.

Subramanian, S, DR Denver, CD Millar, T Heupink, A Aschrafi, SD Emslie, C Baroni, DM Lambert. 2009. High mitogenomic evolutionary rates and time dependency. Trends in Genetics 25:482-486.

Thorne, JL, H Kishino, IS Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. Molecular Biology and Evolution 15:1647-1657.

Wu, CI, WH Li. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proceedings of the National Academy of Sciences of the United States of America 82:1741-1745.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. Journal of Molecular Evolution 39:306-314.

Yang, Z. 1996. Among-site rate variation and its impact on phylogenetic analyses. Trends in ecology & evolution 11:367-372.

Yang, Z, B Rannala. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo Method. Molecular Biology and Evolution 14:717-724.

Zuckerkandl, E, LB Pauling. 1962. Molecular disease, evolution, and genetic heterogeneity. In: M Kasha, B Pullman, editors. Horizons in Biochemistry: Academic Press. p. 189-225.

Chapter 2

Ancient DNA from Pleistocene Chinese fossil cave hyenas reconstructs the recent Eurasian history of the spotted hyena, *Crocuta crocuta*

Statement of authorship

Ancient DNA from Pleistocene Chinese fossil cave hyenas reconstructs the recent Eurasian history of the spotted hyena, *Crocuta crocuta*

Gui-Lian Sheng
Performed DNA extractions, co-developed the research concept, co-wrote the manuscript and acted as corresponding author.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .... *Date* 02/04/2012

Julien Soubrier (Candidate)
Performed the phylogenetic analyses, co-developed the figures and participated to the writing of the manuscript.
I hereby certify that the statement of contribution is accurate.

Signed .... *Date* 02.04.12

Jin-Yi Liu
Assisted with taxonomic and paleonotological interpretation.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .... *Date* 4/5/2012

Lars Werdelin
Co-developed the research concept. Provided paleontological and contextual information, and assisted with interpretation of results.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ..... *Date* April 4, 2012

Michael Hofreiter
Co-developed the research concept, and interpretation of results.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed *Date* 04/04/2012

Bastien Llamas
Contributed to data generation and interpretation. Co-wrote the manuscript.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ……. …… *Date* 02 April 2012

Xu-Long Lai
Co-developed the concept and experimental design, assisted with manuscript preparation.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ... …… *Date* *April 5th ,2012……*

Alan Cooper
Co-developed the concept and experimental design, and co-wrote the manuscript.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …. *Date* *……05.04.2012……*

Ancient DNA from Pleistocene Chinese fossil cave hyenas reconstructs the recent Eurasian history of the spotted hyena, *Crocuta crocuta*

Gui-Lian Sheng[1,2], Julien Soubrier[2], Jin-Yi Liu[3], Lars Werdelin[4], Michael Hofreiter[5], Bastien Llamas[2], Xu-Long Lai[1*], Alan Cooper[2*]

[1] State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Wuhan, Hubei, 430074 China

[2] Australian Centre for Ancient DNA, School of Earth & Environmental Sciences, University of Adelaide, Adelaide, SA, 5000 Australia

[3] Key Laboratory of Evolutionary Systematics of Vertebrates, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing, 100044 China

[4] Department of Palaeozoology, Swedish Museum of Natural history, Box 50007, S-104 05, Stockholm, Sweden

[5] Department of Biology, The University of York, Wentworth Way, Heslington, York YO10 5DD, UK

ABSTRACT

The four living hyena species (spotted, brown, striped and aardwolf) are the remnants of a formerly diverse group of more than 80 fossil species with a peak of diversity in the Late Miocene (about 7-8 My). The taxonomy and evolutionary history of the living species remains contentious, but is of importance for conservation studies due to the dramatic loss of both diversity and geographic range of hyena species since the Pleistocene. The fossil history of living hyena species shows a clear African origin, while morphological and ancient DNA data have confirmed that extinct late Pleistocene cave hyenas from Europe and Asia were closely related to living spotted hyenas (*Crocuta crocuta*). This relationship has been used to explain the origins of Eurasian cave hyena populations via multiple migrations out of Africa between 1.3-3.5 My. We use mitochondrial DNA sequences from Pleistocene hyena specimens collected from northern China to examine the origin of Asian populations and temporally calibrate the evolutionary history of spotted hyenas. Phylogenetic analyses confirm the Asian hyena lineage forms the basal branch within spotted hyenas, and support a far more recent evolutionary timescale for spotted hyenas (163-430 ky) when radiocarbon-dated bones are used as internal calibration points. The new timescale is more consistent with the fossil record, and suggests that a widespread Eurasian population was the ancestor to both extinct and living spotted hyena populations, before retreating to Africa in the Late Pleistocene.

INTRODUCTION

The Hyaenidae family arose in Eurasia in the Late Oligocene, about 25 million years ago (My), and reached a peak of diversity during the Late Miocene, about 7-8 My, totalling more than 80 fossil species from Europe, Asia, Africa, and North America (1-4). The number of Hyaenidae species declined from the Late Miocene, eventually leaving only four extant species (spotted hyena, striped hyena, brown hyena, and aardwolf) in Africa, the Middle East, parts of southwest Asia and the Indian subcontinent. Although it is one of the smallest families of the Carnivora, living hyenas occupy a variety of habitat types and fill a surprisingly wide range of ecological niches, and their presence is a useful indicator of ecosystem health (2). Among the four extant species, the spotted hyena (*Crocuta crocuta*) has attracted considerable evolutionary and systematic interest due to a social system similar to that of many primates, in contrast to other gregarious carnivores (2, 5-6). A variety of Late Pliocene and Pleistocene fossil spotted hyena species are recognized, including *Crocuta dietrichi*, *C. eturono*, *C. honanensis*, *C. ultra*, and the so-called 'cave hyenas' in the Far East (*C. crocuta ultima)* and Europe (*C. crocuta spelaea)* (6-8). However, the exact origin and evolutionary history of living spotted hyenas remains unresolved (1, 5, 7-16).

Ancient DNA provides a means to obtain direct evidence of the phylogenetic and population history of spotted hyenas, and has previously been used to show that the extinct Late Pleistocene cave hyenas in western Eurasia were nested within spotted hyenas (16). An evolutionary model developed to explain the origins of recent spotted hyena populations assumed the ancestral population was located in Africa and inferred three separate dispersal events to Eurasia, between 0.35-3.5 My, to explain the genetic divergence between the three mitochondrial clades found in Late Pleistocene spotted hyenas from Europe and the Far East (16). However, the ages of the inferred dispersals (3.5 My in the Far East and 0.35-1.5 My in western Eurasian) are not consistent with fossil data, which suggest much younger colonization dates for *C. crocuta* morphs in these areas. In Asia, the earliest fossil site for *C. crocuta ultima*, Zhoukoudian Location 1 in China, is dated at 400-230 ky (17- 21) while in Europe, *C. crocuta spelaea* is first seen in western Eurasia around 300 ky, although *C. praespelaea* appeared in Spain as early as 780 ky. The discrepancy between the molecular and paleontological timescales suggests that the evolutionary history of spotted hyenas is not fully resolved and further analyses using more ancient DNA sequences from Eurasian specimens are needed to reconcile the molecular and morphological data.

Crocuta have been found in more than 100 Pleistocene sites in 26 Chinese provinces (17, 22-24), and two morphological forms can be distinguished. *C. honanensis* is first detected in the early Pleistocene of Henan Province. *C. crocuta ultima* was the remaining Pleistocene member of the Hyaenidae family in China. Morphological studies suggest that *C. honanensis* is different

species from the living spotted hyena *C. crocuta* (7, 11, 25), in contrast to *C. crocuta ultima* which has been supported as either *C. crocuta* or an Asian sub-species *C. crocuta ultima* (7, 17, 23). *C. crocuta ultima* became extinct around the end of the Pleistocene, i.e, ~11,500 years B.P. (9, 17, 22-23), or potentially as recently as ~7,800 years B.P. (24). The late survival makes it possible for ancient DNA studies to provide molecular data to explore the evolutionary history of spotted hyenas in East Asia.

In this study, we generated mitochondrial cytochrome *b* (cyt *b*) sequences from northern Chinese Late Pleistocene *C. crocuta ultima* specimens and used radiocarbon-dates from cave hyena bones to calibrate the origin of spotted hyenas in Asia, and provide insights into the population history of spotted hyenas in Europe and Africa.

METHOD

Samples

We collected samples of Late Pleistocene fossil hyenas from Lingxian Cave, Qinghangdao City in Hebei Province; Tonghe Bridge, Zhaodong County in Heilongjiang Province; and Da'an Cave, Tonghua County in Jilin Province, which are all in northern China (Figure 1). The Lingxian Cave samples were associated with a faunal assemblage typical for the "*Crocuta-Cervus* Fauna" in Northern China such as deer (*Cervus* sp.), wolf (*Canis Lupus),* fox (*Vulpus* sp*.)*, woolly rhino (*Coelodonta* sp.), and badger (*Meles* sp.) (Liu, unpublished data). In contrast, the Tonghe Bridge specimen was associated with bison (*Bison* sp.), horse (*Equus* sp), mammoth (*Mammuthus* sp.) and woolly rhino (*Coelodonta* sp.), all of which belong to the "*Coelodonta-Mammuthus* Fauna" in Northeastern China (37). Lastly, the Da'an Cave samples were associated with bear (*Ursus* sp.), deer (*Cervus* (*Pseudaxis) grayi* Zdansky), horse (*Elaphus* sp.), and donkey (*Equus* spp.), badger (*Meles* sp.), rhino (*Dicerorhinus*), and antelope (*Gazella* sp.) (Chen, unpublished data).

Samples in Lingxian Cave were estimated to be greater than 50 ky B.P. through stratigraphic information, while a horse bone (*Equus sp*.) associated with the Tonghe Bridge specimens has been AMS-dated at 34.9 ky B.P. (38). A deer bone (*Cervus sp*.) associated with the Da'an Cave samples was AMS-dated at the Quaternary Geology & Archaeological Chronology Laboratory at Peking University, and was dated at 34.47±0.37 ky B.P. Prior to DNA analyses, the samples had been stored in closed, dry containers at room temperature for several years without any chemical treatment.

Ancient DNA Extraction and Amplification

Ancient DNA extractions and PCR reactions were set up in Wuhan in a laboratory dedicated to ancient DNA research, in a building physically separated from post-PCR facilities. Ancient DNA was extracted from 270-345mg of bone/teeth powder following the methods of Rohland & Hofreiter (39). Extract and PCR blanks were performed throughout all experiments to monitor contamination. We attempted to generate a 713 bp long fragment of the mitochondrial cyt *b* gene, using four overlapping primer pairs from Rohland et al. (16) and five newly designed overlapping primer pairs specifically for this study (Table S1 and Figure S1). Amplifications were performed in 20 µL volumes using a two-step multiplex approach (40). Reagent concentrations and cycling conditions were the same as described in Römpler et al. (40). The nine primer pairs were separated into two non-overlapping sets (indicated in Table S1) and amplified using multiplex PCR (first step), before each primer pair was used individually in singleplex PCRs with the corresponding multiplex PCR product as a template (second step). The annealing temperatures were set at 52°C in both steps.

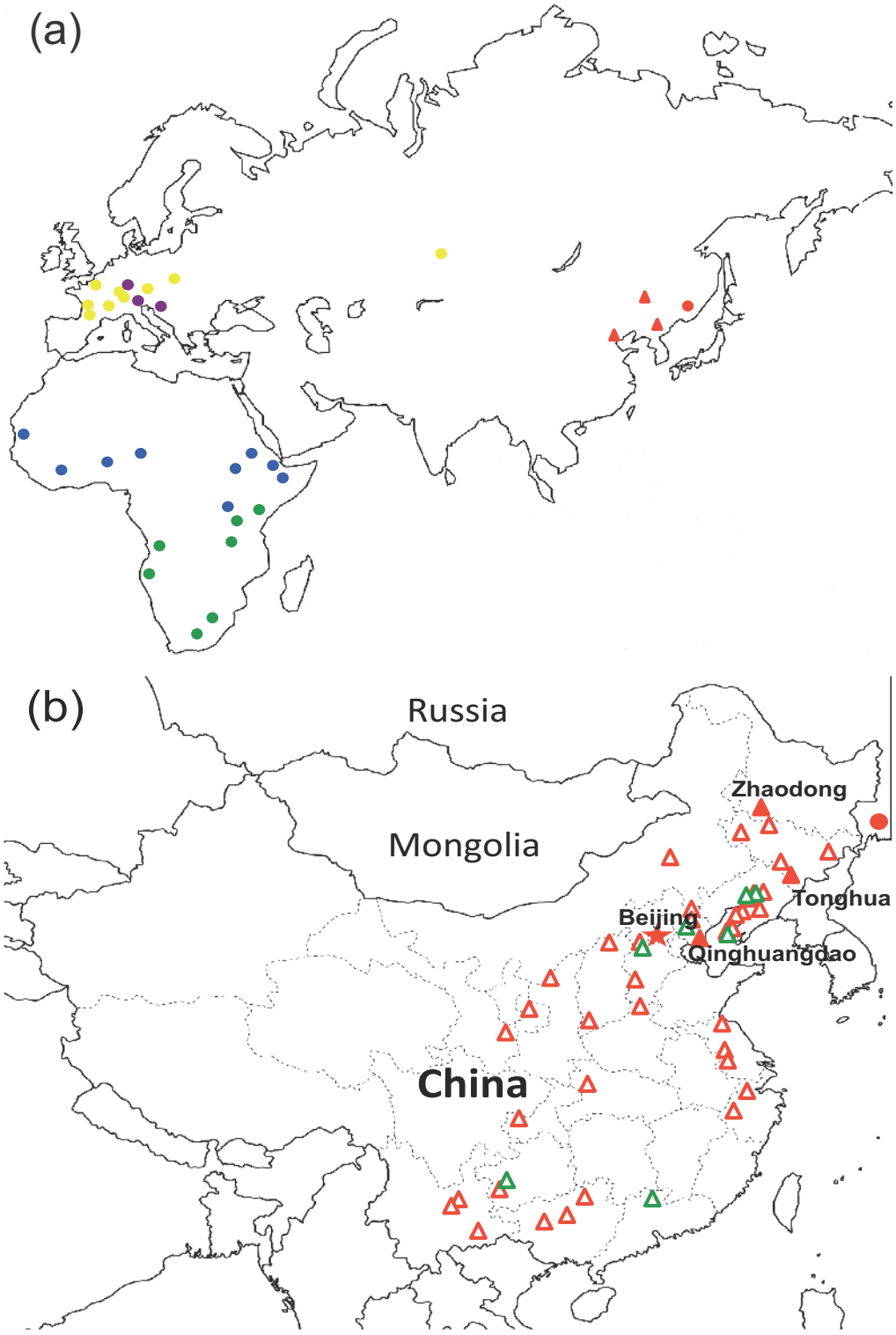


Figure 1. **Phylogeographic distribution of mitochondrial cyt *b* sequences from fossil and living spotted hyenas**. (a) Pleistocene and recent Eurasian and African distribution. Samples in this study are indicated by red filled triangles; samples reported in Rohland et al. (14) are indicated by color-filled circles. (b) Enlarged view of Far East Asia, showing Pleistocene hyena fossil sites in China. Empty red triangles: Late Pleistocene *C. crocuta ultima* hyenas; green triangles: Mid Pleistocene *C. crocuta ultima* morph hyenas.

PCR products were purified using the QIAEX II Gel Extraction Kit (Qiagen, Germany) and cloned into the pMD18-T vector (Takara, Japan) following the supplier's instructions. The recombinant plasmids were transformed into competent E. coli DH5α. White transformants obtained from LB plates containing Amp (0.1mg/mL), X-Gal (0.04 mg/mL) and IPTG (0.024 mg/mL) were screened by PCR with the M13 primer pair. For each fragment, a minimum of eight clones, four from each of two independent primary amplifications, were sequenced at Shanghai Sangon Ltd. Company using an ABI 3700 sequencer following manufacturers' instructions. When consistent differences were found between the first and the second PCR products, due to sequence errors resulting from template damage, a third amplification was performed in 14 out of 27 fragments (shown in Table S1) to determine which sequence was reproducible (41).

Sequence replication

For two of the specimens from Da'an Cave in Jilin Province (DARD-1 and DARD-3), three fragments of cyt *b* (indicated in Table S1, 279bp in total) were independently replicated at ACAD. Teeth samples were extracted using a Qiagen Blood & Tissue Kit (Valencia, California) using a modified protocol with EDTA and Proteinase K added (Thomson et al. in preparation). PCR amplifications were carried out in 25μL volumes, using 1x of PCR buffer, 2.5 mM of MgSO4, 1 mg/mL of rabbit serum albumin (RSA), 0.2 M of each primer, 0.25 mM of dNTPs, 1 U of Platinum HiFi Taq (Invitrogen) and 2 μL of ancient DNA extract. Cycling conditions were: 94°C for 2 min; 50 cycles of 94°C for 15 s, 54°C for 30 s, and 68°C for 20 s; 68°C for 10 min. PCR products were visualized under UV light on a 3.5% agarose gel stained with ethidium bromide. Successful amplifications were purified using Ampure (Agencourt) according to manufacturer's instructions and both strands were sequenced directly using Big Dye chemistry and an ABI 3130XL Genetic Analyzer (Applied Biosystems).

Alignment and Phylogenetic Analyses

Sequence alignments were carried out using the software package Geneious Pro. 5.3.4 (42) and the assemblies were checked manually. A total number of 40 cyt *b* sequences were retrieved from GenBank: 14 fossil spotted hyena, 19 extant spotted hyena, 2 striped hyena (*Hyaena hyaena*), 2 brown hyena (*Parahyaena brunnea*), and 3 aardwolf (*Proteles cristatus*) (Table S2). Three datasets were used to initiate different analyses: i) a first dataset including 3 Chinese Pleistocene cave hyenas, 4 extant spotted hyenas, 2 striped hyenas, 2 brown hyenas, and 3 aardwolf for which 713bp of cyt *b* were available, to establish the phylogenetic position of the Chinese cave hyena in the Hyaenidae family; ii) a second dataset of 366bp cyt *b*, which included 3 Chinese Pleistocene cave hyenas, 14 fossil spotted hyenas, and 15 modern samples where the

specific sampling locations could be traced, to permit network analyses to investigate the relationships among haplotypes of spotted hyenas; and iii) a third dataset of 366bp that included 3 Chinese Pleistocene cave hyenas, 10 fossil spotted hyenas that have radiocarbon dates, and 19 modern samples, which was used for phylogenetic analyses and molecular dating between spotted hyena populations.

We first compared the sequence variation among 7 sequences where at least 713bp of cyt *b* were available, using a cyt *b* sequence of the extant spotted hyena (Table S2, GenBank #AY048786) as reference. To establish the position of the Chinese cave hyena in the Hyaenidae family, a phylogenetic analysis of the 713bp alignment was performed using the program PhyML v3 (43). The best of NNI and SPR moves were used for the tree topology search, and the substitution model TN+I was selected through comparison of Bayesian Information Criterion scores in ModelGenerator v0.85 (44). To investigate the relationships among haplotypes of spotted hyenas, we conducted a median joining network from the second dataset using the program Network v4.6.0.0 (45).

We then performed Bayesian analyses on the third dataset using BEAST v1.6.1 (46) to address the phylogenetic relationships and dates between the cave hyenas from China, *C. spelaea* specimens from other locations in Eurasia, and living spotted hyenas in Africa. The substitution model HKY+G was selected through comparison of Bayesian Information Criterion scores in ModelGenerator v0.85 (47). The MCMC analyses were run for 100,000,000 iterations, with posterior samples drawn every 10,000 steps. Results were checked in Tracer v1.5 after 10% burn-in, and all parameters showed sufficient sampling, indicated by effective sample sizes above 200 (48). A strict molecular clock was used and a constant population size was set as tree prior. Additional analyses using an uncorrelated lognormal relaxed-clock or a Random Local Clock were used to account for potential rate variation, but neither could reject the strict clock assumption, and both led to similar results with no significant differences between divergence dates according to the 95% HPD. A Bayesian skyline analysis was performed to account for potential demographic variation, but after a Bayes factor comparison (49), this model did not show any significant improvement over the assumption of a constant population size. To deal with the issue of time dependence of molecular rates (32, 50), and the effect of deep calibrations on intraspecific dating (51), two sets of dates were used as calibration points to calculate the most recent common ancestor (MRCA) of fossil and extant spotted hyenas: i) the radiocarbon dates associated with the ancient sequences; ii) a log normal distribution with a minimum bound of 9 My and 95% of the prior probabilities younger than 9.5 My for the divergence between Spotted hyenas and Striped/Brown hyenas (here represented by a Striped specimen), in addition to the tip calibrations from the dates of ancient samples.

To check if the signal from the radiocarbon dates associated with the ancient sequences was sufficient to calibrate the spotted hyena phylogeny, a 'date randomization test' was performed (32). All dates associated with the sequences were randomized before the phylogenetic analysis in BEAST was replicated as described above. If the structure and spread of the ancient sequences in the tree show enough temporal information to calibrate the analysis, the inferred mean rate calculated using the correct association date/sequence should be significantly different from the 95% HPD rate estimates calculated from the randomized data set. A comparison of resulting rate estimates from ten replicates and the non-randomized data are shown in Figure S3.

RESULTS

Sequence variation in Pleistocene fossil and extant spotted hyenas

Ten Late Pleistocene hyena specimens from northern China (Figure 1) were analyzed using appropriate ancient DNA techniques (27) including independent extraction and replication, and multiple sequencing reactions (Table S1). Nine overlapping fragments (size range is 127-171 bp including primers) of a 713bp-long region of the mitochondrial cyt *b* gene were amplified from three specimens from Da'an Cave at Tonghua County in Jilin Province using multiplexed PCR (Table S1 and Figure S1). Only one 100bp-long fragment could be amplified for the Tonghe Bridge sample collected from Zhaodong County, Heilongjiang Province, while no fragments could be recovered for six Lingxian Cave samples collected from Qinhuangdao City, Hebei Province, despite multiple attempts. As a result, we considered only the Da'an Cave samples for further phylogenetic analyses.

The three new Chinese *C. crocuta ultima* sequences (DARD-1, DARD-2, and DARD-3) differed from each other at 1-2 positions across the full 713 bp sequence, and from a reference modern spotted hyena sequence (GenBank #AY048786) at 36, 37, and 35 polymorphic sites (~5%) respectively. In contrast, the other three available homologous sequences from GenBank for living spotted hyena have only 18, 18, and 24 polymorphic sites (~3%) across the same fragment (Table S3). The short 100bp-long sequence obtained from the Tonghe Bridge specimen matched the Da'an Cave sequences, which share the same haplotype as the only other Far East Asian Late Pleistocene cave hyena sequenced to date (DQ157555, Pacific coast of Russia), for the shorter 366 bp fragment of cyt *b* generated in a previous study of cave hyena from Europe and Russia (16).

Phylogenetic network and trees

A network analysis of the three new full-length Chinese ancient DNA sequences with 29 published sequences representing both extinct cave hyenas and living spotted hyenas retrieved the same network profile as in Rohland et al. (16), distributed into four haplogroups: the extinct western Eurasian and extant northern African lineage A; the extinct western European lineage B; the extant southern African lineage C, and the extinct Far East Asian lineage D (Figure 2). Lineage D includes the three new Chinese sequences from this study, plus the previously sequenced Russian Pacific Coast specimen.

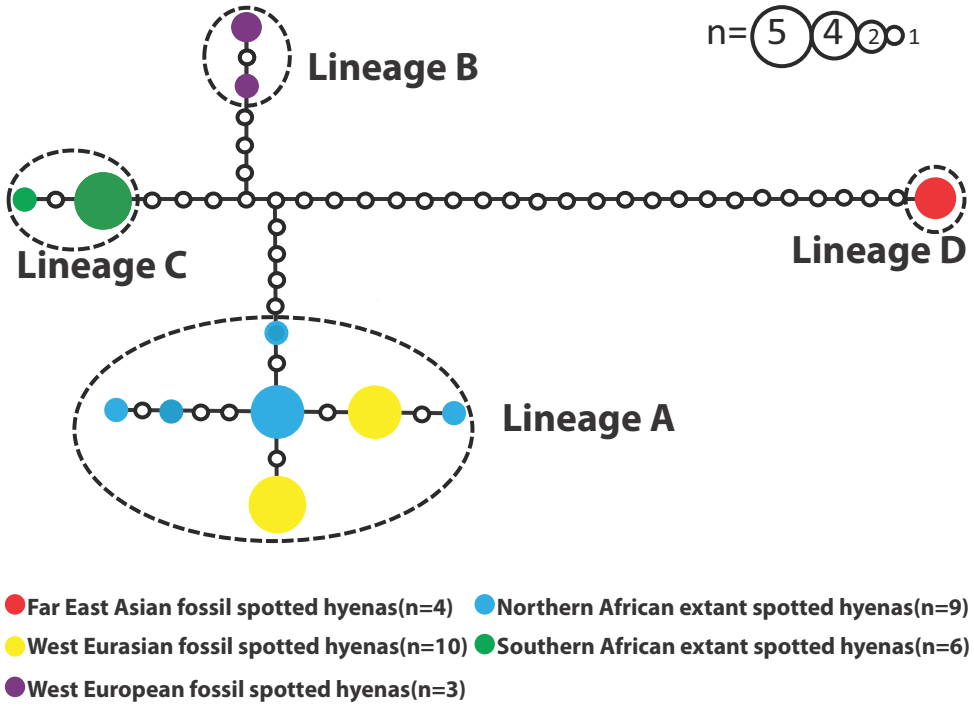


Figure 2 **Median joining network of late Pleistocene and living spotted hyenas.** Analyses based on 366 bp mitochondrial cyt *b* sequences of spotted hyenas, using only ancient and modern specimens with precise geographic localities from Table S1.

The phylogenetic position of the Chinese cave hyenas (*C. crocuta ultima*) in the Hyaenidae was analysed using the 713 bp dataset, and homologous sequences of the four extant species in Hyaenidae. Bayesian analyses showed that the Chinese specimens clustered with the living spotted hyenas and formed a sister group to the branch containing striped hyena and brown hyena (bootstrap value of 98%), with aardwolf in the basal position in the phylogenetic family tree (Figure 3a), as previously suggested by both morphological analyses (7) and molecular studies based on nuclear gene data with bootstrap value of 96% in MP and 84% in ML phylograms (15).

Bayesian analyses were used to investigate the phylogenetic relationships and divergence dates of the various cave and spotted hyena populations. The dataset was limited to a 366bp region to facilitate comparison with previously published sequences from radiocarbon-dated ancient cave hyena bones, and aligned with both modern spotted hyena sequences and a striped hyena sequence as an outgroup.

The sequences from Far East Asia (China and Russia) formed a long basal branch within *C. crocuta,* supported by a bootstrap value of 100%, followed by lineage A (Europe and Africa), separating from the ancestor of lineages B and C (Figures 3b and S2). The estimated dates for the divergence events varied according to the calibration points used. A set of older dates similar to previous estimates (16) was obtained when the basal divergence between the *Crocuta* lineage and *Hyaena/Parahyaena* (striped/brown hyenas) was set at 9-9.5 My, and used in combination with the radiocarbon-dated cave hyena sequences which act as tip dates on the phylogenetic tree (Figure S2). The divergence of lineage D was estimated at 3.22 My (confidence interval [CI]: 1.2 – 5.1 My), while lineage A separated from the ancestor of lineage B and C at 1.52 My (CI: 0.58 – 2.5 My). The final divergence event saw lineage B separate from lineage C at 1.15 My (CI: 0.35-2 My). However, a set of much younger dates was retrieved when the radiocarbon-dated ancient DNA sequences were used as the sole calibration method. This suggests that lineage D diverged from the ancestral spotted hyena population about 430 ky (CI: 134 - 860 ky), followed by lineage A around 224 ky (CI: 84 - 435 ky). Lastly, lineage B and C diverged around 163 ky (CI: 65-315 ky). To test whether the ancient DNA tip dates contained sufficient information to accurately calibrate molecular rate calculations, the dates were randomized with respect to the sequences and re-analyzed. The 95% High Posterior Density (HPD) of the mean rate calculated from the randomized analyses showed no overlap with the rate estimated without randomization, indicating that the radiocarbon dates associated with the ancient samples contain enough temporal information to calibrate the phylogeny (Figure S3).

DISCUSSION

Authenticity of Pleistocene sequences

Ancient specimens contain only trace amounts of highly-fragmented DNA molecules and it is important to demonstrate that the ancient DNA sequences have not been affected by contamination or DNA damage (27-29). In this case, nine overlapping PCRs were used to build the 713bp contigs for each specimen and the sequences were obtained from multiple extractions, amplifications and cloning with clean control reactions. Most importantly, the same sequences were obtained when the experiments were independently replicated at the State Key Laboratory for Biogeology and Environmental Geology at the Chinese University of Geosciences, Wuhan and the Australian Centre for Ancient DNA (ACAD) at the University of Adelaide, Australia.

Previous molecular studies (16) have suggested that neither species nor subspecies status was appropriate for the Late Pleistocene Eurasian cave hyenas (*C. crocuta spelaea*), lineages A and B in Figure 3b, as these are nested within the variation of living spotted hyena populations. In contrast, the Far East Asian cave hyena specimens (currently *C. crocuta ultima*) form the basal lineage (D) with a notably long branch. However, this is currently insufficient to draw taxonomic conclusions, especially in the absence of morphological characteristics differentiating the Far East Asian specimens from the western European specimens, while both of them are distinct from the extant *C. crocuta*.

The older estimates for the divergence events within the spotted hyena phylogeny (1.15-3.22 My) obtained using the 9-9.5 My basal calibration, along with the radiocarbon tip calibrations (Figure S2), were similar to a previous study which just used a 10 My point calibration for the divergence of *Crocuta* and *Hyaena/Parahyaena* (16). The latter study estimated divergences of lineages D, A, and B/C at 3.48, 1.46, and 1.26 My respectively, and explained the Pleistocene phylogeography of spotted hyena by suggesting that these dates represented three migrations out of Africa to Eurasia. However, there are several reasons to question this scenario. Firstly, the estimated dates are not consistent with the paleontological data since *Crocuta* fossils from the Late Pliocene or early Pleistocene do not appear to be *C. crocuta* (8, 18). For example, the 3.4 My-old *C. eturono* in west Turkana, 2.3 My-old *C. honanensis* in China, and 1.6-1.2 My-old *C. ultra* in Israel and Levant, are all morphologically distinguishable from living *C. crocuta* (8-9, 30). These fossils are more likely to record the geographical appearance of the genus *Crocuta*, rather than the species *C. crocuta,* in these areas. Secondly, the phylogeographic distribution does not easily fit an Out-of-Africa dispersal model. For example, several steps are required for a genetic connection between the western European lineage B and the southern African lineage C without a contribution from the geographically intermediate

northern African population, while the absence of lineages B and D in Africa requires multiple extinction events to have occurred within the source population.

When the phylogenetic analyses were repeated using just the radiocarbon dates associated with the ancient sequences as calibration points, the topology remained the same, but the estimated divergence dates were almost an order of magnitude younger. Lineages D, A, and B/C were estimated to diverge around 430 ky, 224 ky, and 163 ky (Figure 3b). These younger dates are a far better fit for the paleontological record of *C. crocuta* in both Asia and Europe, with the earliest *C. crocuta ultima* in China detected around 400-230 ky (17, 19, 21), while the cave hyena appeared in western Eurasia sometime after 300 ky. To investigate whether the structure and distribution of the dated sequences within the topology was likely to contain sufficient evolutionary information to generate a meaningful rate estimate the analysis was repeated with the tip dates randomized amongst the ancient sequences. The inferred mean rate was significantly different from the rate calculated when the correct dates were associated with the sequences (Figure S3), indicating there was useful evolutionary information within the temporally distributed sequences.

The much older results generated by the basal 9-9.5 My calibration point are likely to relate to the temporal dependency of molecular rate estimation (31-33). In phylogenetic analyses using hyena outgroups, the molecular rates are measured over a long timescale i.e., 9.5 or 10 My calibration, between distantly related species (spotted hyena and striped/brown hyenas). Under these conditions, the rate estimates approach the substitution rate. In contrast, rate estimates calibrated with just the radiocarbon tip dates cover a much shorter period of evolution and record both intra-population variation (mutation events) as well as substitution events, and are generally considerably faster (32). It is expected that the tip date calibrated rates are more likely to be appropriate for dating recent evolutionary events such as those within the Late Pleistocene, and along with the much better match to the fossil record, suggest these rates should be used to re-assess the recent evolutionary history of spotted hyenas.

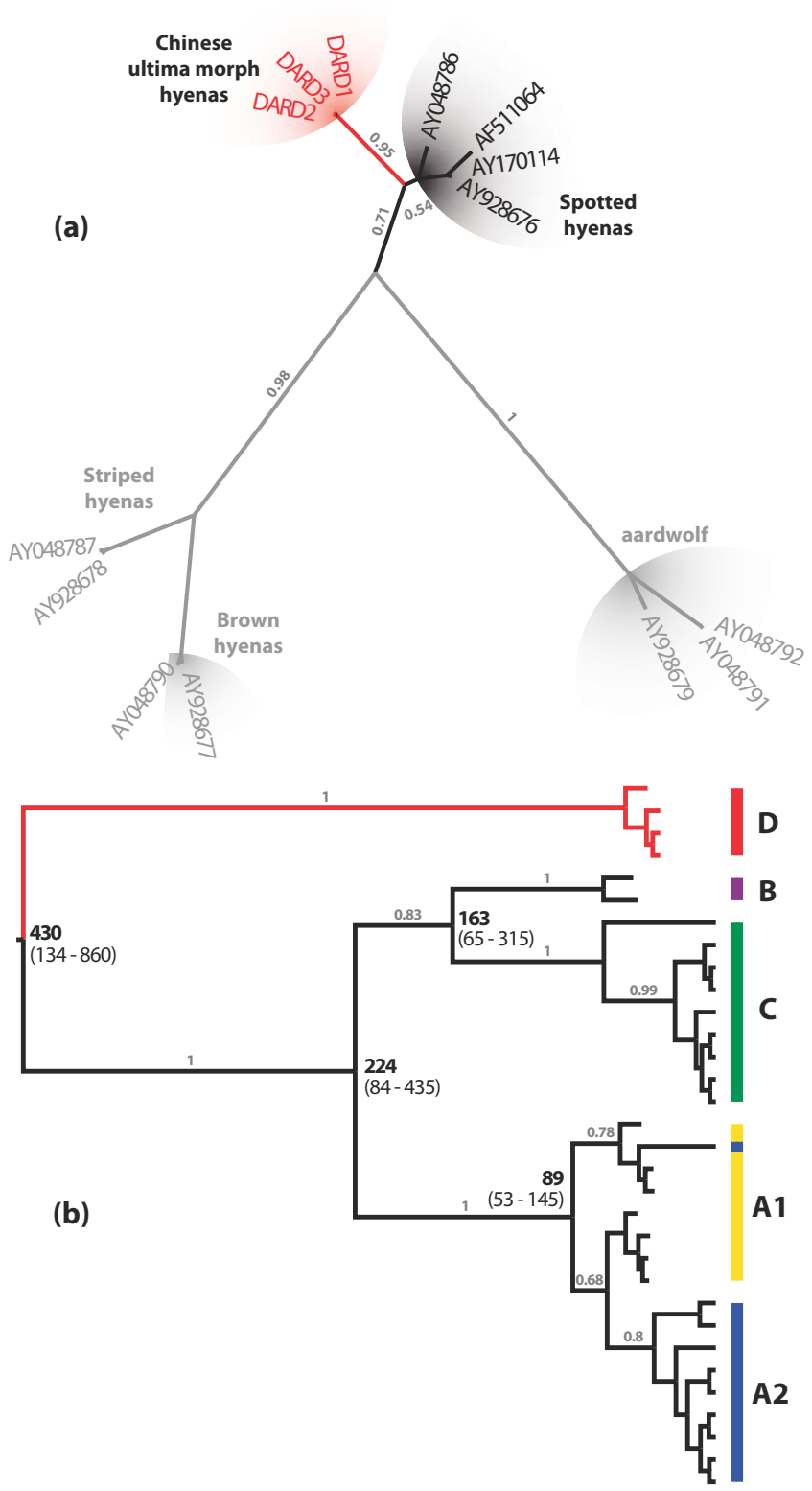


Figure. 3. **Phylogenetic analyses**. **(a)** Maximum likelihood tree of the Hyaenidae based on the 713bp cyt *b* dataset, indicating the position of the Chinese cave hyenas. **(b)** Maximum clade credibility tree estimated using Bayesian analysis of fossil and living spotted hyenas, based on a 366bp dataset and calibrated with the radiocarbon dates of all ancient samples as tip dates. The colors correspond to the different geographic areas shown in Figure 1. Molecular rate calculations show that the sequences from Far East Asia form the basal branch within *C. crocuta,* and diverged about 430 ky, followed by the split between lineage A and the ancestor of lineage B and C around 224 ky. Lineage B and C diverged around 163 ky, while the western Eurasian lineage (A1) and the southern African lineage (A2) separated about 89 ky.

Recent central Eurasian evolutionary history of spotted hyenas

The topology of the phylogenetic trees, and the fossil and molecular dates suggest an alternative scenario for the recent evolutionary history of spotted hyena, based around the fragmentation of an ancestral population located in the large Eurasian steppe ecosystem. The steppe ecosystem stretched from central to western Eurasia in the mid-Early Pleistocene and provided a habitat for both ungulates and other carnivores, such as cave lions, which are known to have existed across this region (33-35). An ancestral Eurasian hyena population is proposed to have contained ancestral *A**, *B**, *C** and *D** lineages, amongst others, before being fragmented by changing environmental conditions in the Late Pleistocene (Figure 4). The far eastern portion of the population is suggested to have become isolated sometime after 430 ky, separating the ancestor of lineage D. Some time after 90 ky, when lineages A1 and A2 are estimated to have diverged, the western Eurasian population (lineages A1 and B) also became isolated. The remnant ancestral Eurasian population then retreated (back) into Africa where lineage A and C eventually became sorted into northern and southern populations (Figure 4). The low diversity within each haplogroup in *C. crocuta* indicates that the spotted hyena populations have been subjected to population bottlenecks in the past, which likely resulted in the loss of lineages from different areas. This could explain the phylogeographic structure, such as the fixation of lineage D in eastern populations and its absence elsewhere, or the lack of lineage C in Europe.

Modern African hyena populations (both lineages A2 and C) are morphologically distinct from fossil populations (7). This distinctiveness appears very late in the fossil record and is observed first in the Middle East in Israel (30; marked with # on Figure 4d). The exact timing of the appearance of the extant morph is, however, unclear, since the record of Late Pleistocene *Crocuta* is poor outside Europe and China, but must have been later than 1 My, since an archaic morph of *Crocuta* is known from Olorgesailie, Kenya, in sediments dated 0.99 My (36). Such a late morphological shift would be compatible with the remnant Eurasian population entering, or being confined to, Africa. Consequently, the new model provides a good fit for both the timing and morphological shifts observed within the paleontological record. In contrast, such a morphological transition is not easily reconciled with the previous Out of Africa model.

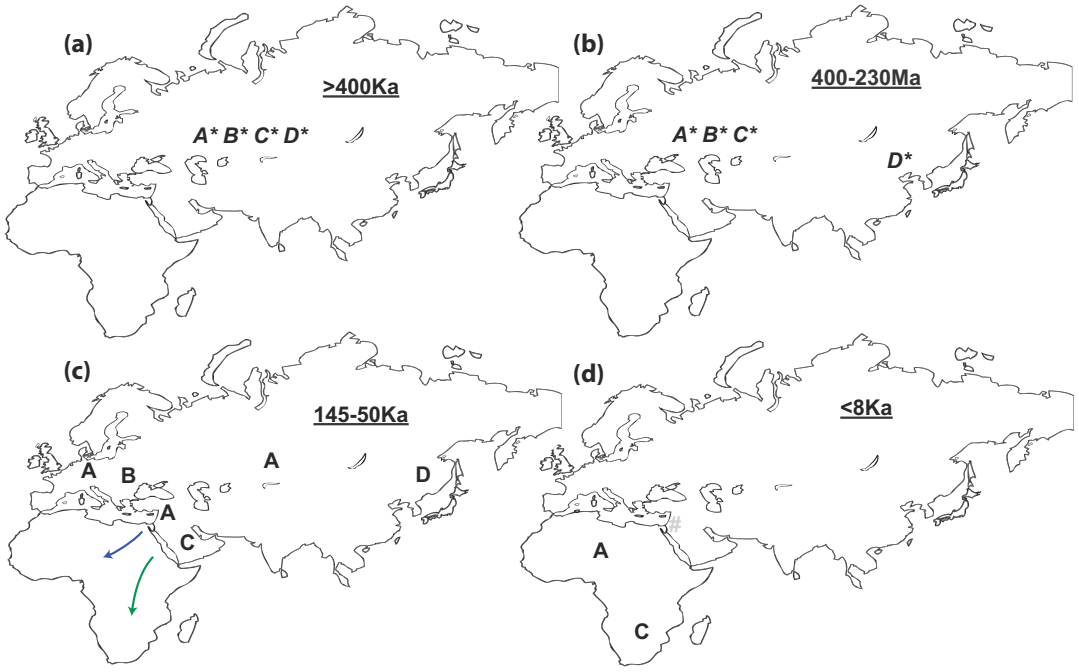


Figure 4. **Revised model of the recent evolutionary history of spotted hyenas**. **(a)** A large ancestral population is proposed to have existed across the steppe habitat of central Eurasia in the mid-Early Pleistocene, containing ancestral lineages of A, B, C, and D (indicated with italics and asterisks); **(b)** The ancestral population began to fragment in the late middle Pleistocene, potentially in response to environmental changes. At some point after 430 ky, the ancestral lineage *D** was isolated in Far East Asia; **(c)** Lineage A1 and B were isolated in western Eurasia at some point after the 90 ky estimated separation of the A1 and A2 lineages, and the remnant Eurasian population retreated into the middle East/Africa; **(d)** The African population was subsequently sorted into northern (A2) and southern (C) lineages while lineages A, B, and D became extinct in Eurasia during the Late Pleistocene.
Grey areas in (a) and (b) show the likely distribution of the ancestral population for different lineages. Colored areas in (c) and (d) show the real distribution of spotted hyenas, which is identified by either the fossil records or the extant populations. Dates in (a), (b), and (d) are based on fossil dating that the oldest fossil *Crocuta crocuta* in China was dated as 400 ky or 230 ky; and the ultima morph became extinct in China about 8 ky. Dates in (c) come from molecular dating and the radiocarbon date: 50 ky is the oldest radiocarbon date of the ancient sample in our dataset; in the phylogenetic tree, the latest separation event took place around 53-145 ky in lineage A, between the western Eurasian lineage (A1) and the northern African lineage (A2), which means that as far as 145 ky, the four lineages did exist in the shadow areas. The # sign represents the geographical location of the oldest *Crocuta* remains morphologically distinct from fossil lineages.

Acknowledgements

We are grateful to Jian Yi, Jun-Xia Yuan in BGEG for their technical support. We thank Jeremy Austin, Wolfgang Haak, and members in ACAD for their helpful comments on the manuscript. This research was supported by 973 Program (No.2011CB808800), the Natural Science Foundation of China (Nos. 40902008, 40972013), and the Australian Research Council.

References

1. Werdelin L, Solounias N (1991) The hyaenidae: taxonomy, systematics and evolution. *Fossils Strata*, **30**,1-104
2. Turner A, Antón M, Werdelin L (2008) Taxonomy and evolutionary patterns in the fossil Hyaenidae of Europe. *Geobiology*, **41**: 677-687
3. Werdelin L, Yamaguchi N, Johnson W E, O'Brien SJ (2010) Felid phylogeny and evolution. In: Macdonald D M, Loveridge A (eds.) The Biology and Conservation of Wild Felids, pp. 59-82. Oxford University Press
4. Bonis L De（1973）Contribution à l'étude des mammifères de l'Aquitanien de l'Agenais: rongeurs, carnivores, périssodactyles. Mémoires du Muséum national d'Histoire naturelle, Nouvelle Série (C) 28: 192P, Paris
5. Watts HE, Holekamp KE (2007) Hyena societies. *Current Biology*, **17**(16): 657-660
6. Diedrich C (2008) The holotypes of the upper Pleistocene *Crocuta crocuta spelaea* and *Panthera leo spelaea* of the zoolithen cave hyena den and their palaeo- ecological interpreitation. *Zoological Journal of the Linnean Society*, **154**, 822-831
7. Kurtén B(1956) The status and affinities of *Hyaena sinensis* Owen and *Hyaena ultima* Matsumoto. *Am. Mus. Novit.*, **1764**, 1-48
8. Werdelin L, Lewis ME (2008) New species of *Crocuta* from the early Pliocene of Kenya, with an overview of early Pliocene hyenas of eastern Africa. Journal of Vertebrate Paleontology, 28(4): 1162-1170
9. Kurtén B(1968) Pleistocene mammals of Europe. Weidenfeld and Nicholson, London.
10. Turner A (1984) The interpretation of variation in fossil specimens of spotted hyena (*Crocuta crocuta* Erxleben, 1777) from Sterkfontein valley sites (mammalian: carnivora). *Ann. Transvaal Mus.*, **33**: 399-418
11. Huang WP (1989) Taxonomy of the hyaenidae (*hyaena* and *crocuta*) of the pleistocene in China. *Vertebrata Palasiatica*, **27**(3),197-204 (in Chinese with English abstract)
12. Markova AK, Smirnov NG, Kozharinov AV, Kazantseva NE, Simakova AN, Kitaev LM (1995) Late Pleistocene distribution and diversity of mammals in northern Eurasia (PALEOFAUNA Database). *Paleontol. Evol.*, **28-29**: 5-134
13. Jenks SM, Werdelin L (1998) Taxonomy and systemetics of living hyaenas (Family Hyaenidae). In: Mills M G L, Hofer H (Compilers). Hyaenas: Status survey and conservation action plan. IUCN/ SSC Hyaena Specialist Group, IUCN Gland, Switzerland, pp. 8-17
14. Werdelin L (1999) Studies of fossil hyaenas: affinities of *Lycyaenops rhomboideae* Kretzoi from Pextlörinc, Hungary. *Zoological Journal of the Linnean Society*, **126**:307-317

15. Koepfli KP, Jenks SM, Eizirik E, Zahirpour T, Valkenburgh BV, Wayne RK (2006) Molecular systematics of the hyaenidae: relationships of a relictual lineage resolved by a molecular supermatrix. *Molecular Phylogenetics and Evolution*, **38**,603-620
16. Rohland N, Pollack J L, Nagel D, Beauval C, Airvaux J, Paabo S, Hofreiter M (2005) The population history of extant and extinct hyenas. *Molecular Biology and Evolution*, **22**(12): 2435-2443
17. Liu JY (1999) Study of the Late Pleistocene mammalian remains from Miaofengshan, Beijing with discussion of the migration and extinction of *Crocuta crocuta* ultima. *Longgupo Prehistoric Culture*, **1**: 128-139 (In Chinese)
18. Qiu ZX, Deng T, Wang BY (2004) Early Pleistocene mammalian fauna from Longdan, Dongxiang, Gansu, China. *Palaeontologia Sinica*, New Series C, 2004, 27:1-198
19. Hu CK (1985) The history of Mammalian Fauna of Locality 1 of Zhoukoudian and its recent advances. In Wu RK et al., eds: Multi-displinary study of the Peking Man Site at Zhoukoudian. Beijing, Science Press, 107-113 (In Chinese)
20. Zhou C, Lui Z, Wang Y (2000) Climatic cycles investigated by sediment analysis in Peking Man's Cave, Zhoukoudian, China. *Journal of Archaeological Science*, **27**:101-109
21. Shen GJ, Ku TL, Cheng H, Edwards RL, Yuan ZX, Wang Q (2001) High-precision U-series dating of Locality 1 at Zhoukoudian, China. *Journal of Human Evolution*. 41**:**679–688
22. Tong HW (2007) Occurrences of warm-adapted mammals in north China over the Quaternary period and their paleoenvironmental significance. *Science in China (Series D)*, **50**(9), 1327-1340
23. Tseng ZJ, Chang CH (2007) A study of new material of *Crocuta crocuta ultima* (Carnivora: Hyaenidae) from the Quaternary of Taiwan. *Coll. and Res*. **20**: 9-19
24. Ma AC, Tang HL (1992) On discovery and significance of a Holocene Ailuropoda-Stegodon Fauna from Jinhua, Zhejiang. Vertebrate PalAsiatica, 30(4): 295-312 (In Chinese with English summary)
25. Qiu ZX(1987) Die Hyaeniden aus dem Ruscinium und Villafranchium Chinas. Mnchner Geowissenschaftliche Abhandlungen A, **9**: 1-108
26. Zdansky O (1924) Jungtertiäre Carnivoren Chinas. *Palaeontol. Sin. Series C*, **2**(1): 1-149
27. Cooper A, Poinar HN (2001) Ancient DNA: do it right or not at all. *Science*, **289**, 1139
28. Poinar HN (2003) The top 10 list: criteria of authenticity for DNA from ancient and forensic samples. International Congress Series 1239, 575e579

29. Knapp M, Clarke AC, Horsburgh KA, Matisoo-Smith EA (2011) Setting the stage---- building and working in an ancient DNA laboratory. *Annals of Anatomy*, doi:10.1016/j.aanat.2011.03.008
30. Martinez-Navarro B, Belmaker M, Bar-Yosef O (2009) The large carnivores from 'Ubeidiya (early Pleistocene, Israel): biochronological and biogeographical implications. Journal of Human Evolution 56: 514-524
31. Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**(3) 367-380
32. Ho SYW, Lanfear R, Bromham L, Phillips MJ, Soubrier J, Rodrigo AG, Cooper A (2011) Time-dependent rates of molecular evolution. *Molecular Ecology*, **20**: 3087-3101
33. Yamaguchi N, Cooper A, Werdelin L, Macdonald DW (2004) Evolution of the mane and group-living in the lion (*Panthera leo*): a review. *J. Zool. Lond.* **263**:329-342
34. Barnett R, Yamaguchi N, Barnes I, Cooper A (2006) The origin, current diversity and future conservation of the modern lion (*Panthera leo*). *Proc. R. Soc. B*, doi: 10.1098/rspb.2006.3555
35. Barnett R, Shapiro B, Barnes I, Ho SYW, Burger J, Yamaguchi N, Higham TFG, Wheeler HT, Rosendahl W, Sher A V, Sotnikova M, Kuznetsova T, Baryshnikov GF, Martin LD, Harington CR, Nurns JA, Cooper A (2009) Phylogeography of lions (*Panthera leo* ssp.) reveals three distinct taxa and a late Pleistocene reduction in genetic diversity. *Molecular Ecology*, **18**: 1668- 1677
36. Werdelin L, Lewis M E (2005). Plio-Pleistocene Carnivora of eastern Africa: species richness and turnover patterns. *Zoological Journal of the Linnean Society*, **144**: 121-144
37. Chow MC (1959) Age of the mammalian fossil assemblages. In: The Paleomammalogy Group of IVPP ed. Pleistocene Mammalian Fossils from the Northeastern Provinces. *Mem Inst Vertebr Paleontol Paleoanthropol, Acad Sin,* **3**:9-10
38. Weinstock J, Willerslev E, Sher A, Tong W, Ho S W, Rubenstein D, Storer J, Burns J, Martin L, Bravi C, Prieto A, Froese D, Scott E, Lai X L, Cooper A (2005). Evolution, systematics, and phylogeography of Pleistocene horses in the new world: A molecular perspective. *PloS Biology*, **3**(8): 1373–1379
39. Rohland N, Hofreiter M (2007) Ancient DNA extraction from bones and teeth. *Nat Protoc*, **2** (7), 1756-1762
40. Römpler H, Dear PH, Krause J, Meyer M, Rohland N, Schoneberg T, Spriggs H, Stiller M, Hofreiter M (2006) Multiplex amplification of ancient DNA. *Nature*, **1**(2): 720-728
41. Hofreiter M, Jannicke V, Serre D, Haeseler A, Paabo S (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Research*, **29**(23), 4793- 4799

42. Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Heled J, Kearse M, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2010) Geneious v5.3, Available from http://www.geneious.com
43. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate Maximum-Likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*, **59**(3), 307-321
44. Keane TM, Creevey CJ, Pentony MM, Naughton TJ, McInerney JO (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evolutionary Biology*, **6**:29
45. Bandelt HJ, Forster P, Rohl A (1999) Median joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution,* **16**: 37e48
46. Drummond AJ, Rammbaut A (2007) BAEST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214
47. Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009) Bayesian phylogeography finds its roots. *PLoS Computational Biology* **5**, e1000520
48. Rambaut A, Drummond AJ (2007) Tracer v 1.4. Available from: http://beast.bio.ed.ac.uk/Tracer
49. Suchard MA, Weiss RE, Sinsheimer JS (2001) Bayesian selection of continuous-time Markov China Evolutionary models. *Molecular Biology and Evolution*, **18**: 1001-1013
50. Ho SYW, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution*, **22**, 1561-1568
51. Ho SYW, Saarma U, Barnett R, Haile J, Shapiro B (2008) The effect of inappropriate calibration: three case studies in molecular ecology. *PLoS ONE*, **3**, e1615

SUPPLEMENTARY MATERIAL

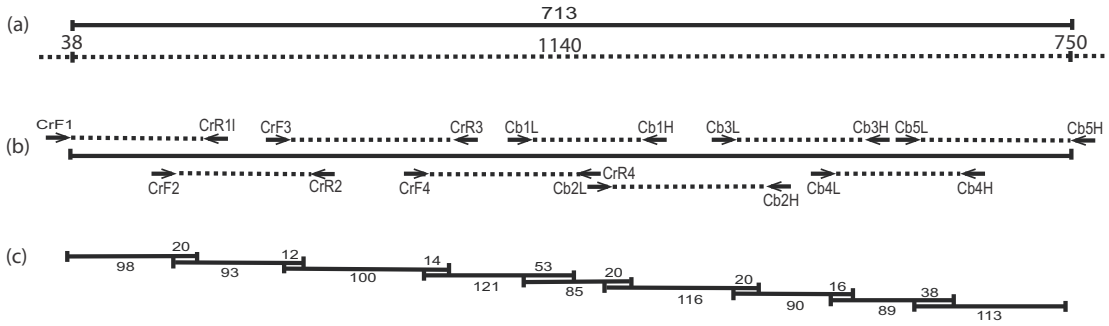


Figure S1. Schematic view of the 713 bp configs of the cyt *b* gene for the Pleistocene samples using nine overlapping PCR fragments. (a) Beginning and ending nucleotide positions of the 713 bp contigs in the 1140 bp complete cyt *b* gene; (b) Primer binding areas of nine primer pairs; (c) Nine overlapping fragments, numbers below fragments show length of the amplification products without primers, numbers above fragments show overlaps between individual fragments.

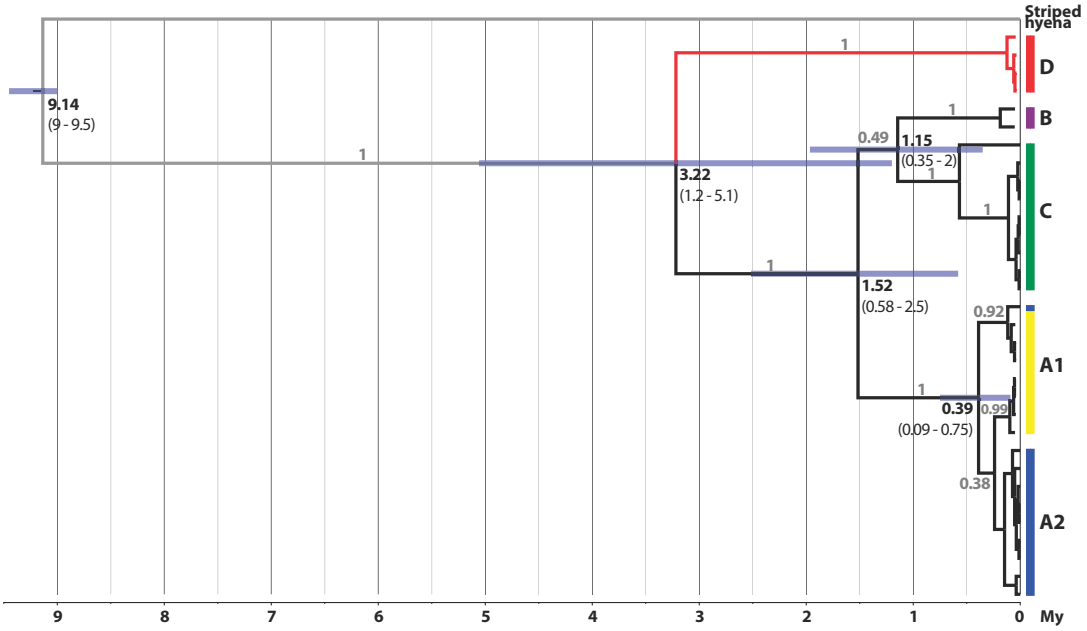


Figure S2. Phylogenetic tree for fossil and extant spotted hyenas from the 366bp dataset, using the striped hyena as an outgroup and calibration of 9-9.5 My for the divergence between *Crocuta* lineage and *Hyaena/Parahyaena* lineage, in addition to the tip calibrations from the dates of ancient samples.

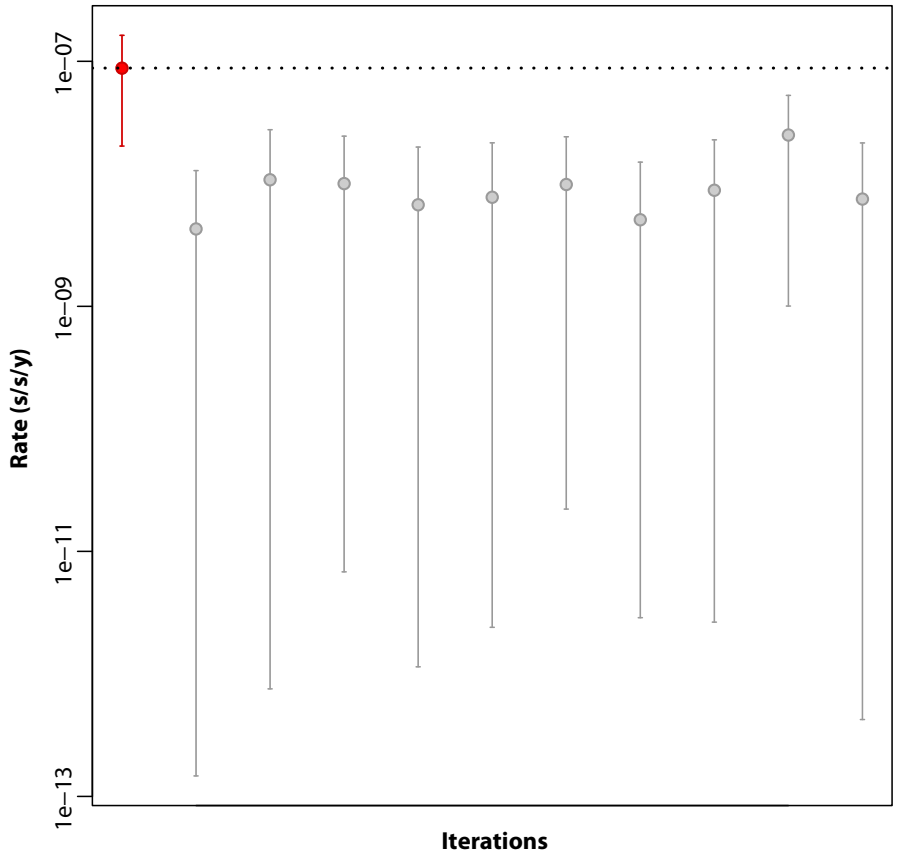


Figure S3. Date Randomization Test. Red circle and dotted line represent the mean rate calculated during the phylogenetic analysis of the 366bp alignment of spotted hyenas using the radiocarbon dates associated with the ancient sequences as calibration. The grey lines represent the 95% HPD of rates calculated for ten replicates of the same analysis with randomized dates. The fact that none of these margins overlap with the original mean rate demonstrates that the radiocarbon dates used for this study is informative enough to calibrate the timed phylogeny.

Table S1. PCR primers for *Crocuta crocuta ultima* mitochondrial cytochrome *b* gene

Primer Name	Primers (3' position in 1140 bp *C. crocuta* mt cyt b gene, AY928676)	Amplicon Size (in/ excluding primers)	Reference	Multiplex set	Third amplification performed		
					DARD-1	DARD-2	DARD-3
CrF1	GAAAATCTCACCCACTCATTAAAA (14)	147/ 98 bp	14	1	✓		
CrR1l	GTATGGCTAGGAATAGACCTGTCAG (160)						
CrF2	GAAATTTCGGGTCACTATTAGGAA (92)	136/ 93 bp	14	2		✓	
CrR2	CCGTAGTTTACGTCTCGGC (227)						
CrF3*	CAACAACCGCCTTCTCATCAG (176)	146/ 100 bp	14	1	✓		✓
CrR3*	GTAAGATCCGTAGTATATTCCTCGG (321)						
CrF4	GAGCTTCCATATTCTTCATCTGTCTA (257)	171/ 121 bp	14	2		✓	
CrR4	CACCTCAGAATGATATTTGGCCTC (427)						
Cb1L*	AGACGTGAAACATCGGAATC (332)	125/ 85 bp	This study	1	✓	✓	✓
Cb1H*	GGCTGATAGGAGGTTGGTAA (456)						
Cb2L	TACCATGAGGCCAAATATC (398)	153/ 116 bp	This study	2	✓		
Cb2H	TAAAGTGGAGGGCGAAGA (550)						
Cb3L*	GAGGAGGATTCTCAGTGGA (494)	128/ 90 bp	This study	1	✓		✓
Cb3H*	GTTGTTGGAGCCTGTTTCG (621)						
Cb4L	CTAGCCCTGGCAACCGTCC (568)	127/ 89 bp	This study	2	✓	✓	
Cb4H	GGCCTAGAATGTCTTTGGT (694)						
Cb5L	AACCCCTCAGGAATAACAT (619)	150/ 113 bp	This study	1	✓		
Cb5H	GTAGTTGTCGGGGTCTCC (768)						

*: primer pairs have been replicated in ACAD

√: a third amplification have been performed to avoid sequence errors due to template damage

Table S2. Datasets used in this study

No. in this study	Species	C14 age (ka)	GenBank Accesion No.	Length	Location	Reference
DARD-1	*Crocuta crocuta ultima*	34.47±0.37	XXX	713 bp	China	This study
DARD-2	*Crocuta crocuta ultima*	34.47±0.37	XXX	713 bp	China	This study
DARD-3	*Crocuta crocuta ultima*	34.47±0.37	XXX	713 bp	China	This study
C.crocuta_Belgium	*Crocuta crocuta spelaea*	N/A	DQ157554	366 bp	Belgium	1
C.crocuta_Russia	*Crocuta crocuta spelaea*	48.65+2.38/-1.84	DQ157555	366 bp	Russia	1
C.crocuta_Austria1	*Crocuta crocuta spelaea*	38.06±0.85	AJ809318	366 bp	Teufelsucke Austria	2
C.crocuta_Austria2	*Crocuta crocuta spelaea*	38.68±0.92	AJ809320	366 bp	Winden, Austria	2
C.crocuta_Czech Rep	*Crocuta crocuta spelaea*	46.0±2.1	AJ809321	366 bp	Czech Rep.	2
C.crocuta_North Sea	*Crocuta crocuta spelaea*	N/A	AJ809323	366 bp	The Netherlands	2
C.crocuta_Romania	*Crocuta crocuta spelaea*	41.8+1.4/-1.2	AJ809324	366 bp	Romania	2
C.crocuta_France	*Crocuta crocuta spelaea*	40.7±0.9	AJ809325	366 bp	France	2
C.crocuta_Ukraine	*Crocuta crocuta spelaea*	41.3±1.2	AJ809326	366 bp	Ukraine	2
C.crocuta_Altai	*Crocuta crocuta spelaea*	42.3+0.94/-0.84	AJ809327	366 bp	Russia	2
C.crocuta_Slovalia	*Crocuta crocuta spelaea*	51.2+4.9/-3.0	AJ809328	366 bp	Slovakia	2
C.crocuta_Germany1	*Crocuta crocuta spelaea*	N/A	AJ809329	366 bp	Germany	2
C.crocuta_Germany2	*Crocuta crocuta spelaea*	N/A	AJ809330	366 bp	Germany	2
C.crocuta_Hungary	*Crocuta crocuta spelaea*	41.8±1.3	AJ809331	366 bp	Hungary	2
C.crocuta_Senegal	*Crocuta crocuta*	modern	DQ157556	366 bp	Senegal	1
C.crocuta_Ethiopial	*Crocuta crocuta*	modern	DQ157557	366 bp	Ethiopial	1
C.crocuta_Cameroon	*Crocuta crocuta*	modern	DQ157558	366 bp	Cameroon	1
C.crocuta_Togo	*Crocuta crocuta*	modern	DQ157559	366 bp	Togo	1
C.crocuta_Tanzania	*Crocuta crocuta*	modern	DQ157560	366 bp	Tanzania	1
C.crocuta_Rwanda	*Crocuta crocuta*	modern	DQ157562	366 bp	NE-Rwanda	1
C.crocuta_Eritrea	*Crocuta crocuta*	modern	DQ157564	366 bp	Eritrea	1
C.crocuta_Sudan	*Crocuta crocuta*	modern	DQ157565	366 bp	Sudan	1
C.crocuta_Zimbabwe	*Crocuta crocuta*	modern	DQ157566	366 bp	Zimbabwe	1
C.crocuta_Namibia	*Crocuta crocuta*	modern	DQ157568	366 bp	Namibia	1
C.crocuta_South Africa	*Crocuta crocuta*	modern	DQ157569	366 bp	South Africa	1
C.crocuta_Angola	*Crocuta crocuta*	modern	DQ157570	366 bp	Angola	1
C.crocuta_Somalia	*Crocuta crocuta*	modern	DQ157571	366 bp	Somalia	1
C.crocuta_Kenya	*Crocuta crocuta*	modern	DQ157572	366 bp	Kenya	1
C.crocuta_Uganda	*Crocuta crocuta*	modern	DQ157574	366 bp	Uganda	1
C.crocuta_zoo1	*Crocuta crocuta*	modern	AY048786	1137 bp	N/A	3
C.crocuta_zoo2	*Crocuta crocuta*	modern	AF511064	1140 bp	N/A	4
C.crocuta_zoo3	*Crocuta crocuta*	modern	AY170114	1140 bp	N/A	5
C.crocuta_zoo4	*Crocuta crocuta*	modern	AY928676	1140 bp	N/A	6
H.hyaena	*Hyaena hyaena*	modern	AY928678	1140 bp	N/A	6
H.hyaena	*Hyaena hyaena*	modern	AY048787	1137bp	N/A	3
P.brunnea	*Parahyaena brunnea*	modern	AY048790	1137bp	N/A	3
P.brunnea	*Parahyaena brunnea*	modern	AY928677	1140bp	N/A	6
P.cristatus	*Proteles cristatus*	modern	AY048791	1137bp	N/A	3
P.cristatus	*Proteles cristatus*	modern	AY048792	1137bp	N/A	3
P.cristatus	*Proteles cristatus*	modern	AY928675	1140bp	N/A	6

Table S3. Variations in the newly obtained ancient sequences compared to living spotted hyenas

	40	46	84	90	96	102	109	124	144	153	179	204	219	243	245	258	261	264	291	321	336	342	358	387	390	393	396	402	426	438	468	474	478	495	503	513	565	574	577	578	648	657	669	705	707	709	712
AY048786	A	A	A	A	T	G	C	G	C	A	C	C	T	C	T	A	T	C	C	T	G	T	C	G	T	T	C	G	T	T	C	C	T	A	T	T	G	C	G	T	C	A	C	G	T	C	T
AY928676	.	.	G	G	.	.	T	A	T	.	.	.	A	.	.	.	.	.	T	C	.	C	.	A	.	.	.	A	.	.	.	T	C	.	.	.	A	.	.	C	.	.	.	.	.	.	T	A
AY170114	.	.	G	G	.	.	T	A	T	.	.	.	A	.	.	.	.	.	T	C	.	C	.	A	.	.	.	A	.	.	.	T	C	.	.	.	A	.	.	C	.	.	.	.	.	.	T	A
AF511064	G	G	.	.	.	.	T	A	T	.	.	.	A	.	.	.	.	.	T	C	.	C	.	A	.	.	.	A	.	C	.	T	C	.	C	.	A	G	.	C	.	.	.	.	.	.	T	A
DARD-1	G	.	.	.	C	A	.	A	.	G	T	T	.	T	C	T	C	T	T	C	A	C	T	.	C	C	T	A	C	.	T	.	C	G	.	C	.	.	A	C	T	G	T	A	C	.	A
DARD-2	G	.	.	.	C	A	.	A	.	G	T	T	.	T	C	T	C	T	T	C	A	C	T	.	C	C	T	A	C	C	T	.	C	G	.	C	.	.	A	C	T	G	T	A	C	.	A
DARD-3	G	.	.	.	C	A	.	A	.	G	T	T	.	T	C	T	C	T	T	C	A	C	T	.	C	C	T	A	C	.	T	.	C	G	.	C	.	.	A	.	T	T	T	A	C	.	A

There are only four long sequences (GenBank Accession Nos: AY048786, AY928786, AY170114, and AF511064) available in GenBank. The large majority of the variable positions in the Pleistocene fossil spotted hyenas were transitions, with only two transversions (A→T at nucleotide position 258 and T→A at nucleotide position 712). Moreover, 19.4%, 10.2%, and 69.4% of the polymorphic sites were found at 1st, 2nd, and 3rd codon positions, respectively.

Supplementary material reference:

1. Rohland N, Pollack J L, Nagel D, Beauval C, Airvaux J, Paabo S, Hofreiter M (2005) The population history of extant and extinct hyenas. Molecular Biology and Evolution, 22(12): 2435-2443
2. Hofreiter M, Serre D, Rohland N, Rabeder G, Nagel D, Conard N, Munzel S, Paabo S (2004) Lack of phylogeography in European mammals before the last glaciation. Proc. Natl. Acad. Sci. USA, 101: 12963-12968
3. Albert R (2001) Gene structure and gene flow in selected populations of spotted hyaena (Crocuta crocuta). Thesis, Freie Universitaet Berlin
4. Gaubert P, Tranier M, Delmas AS, Colyn M, Veron G (2004) First molecular evidence for reassessing phylogenetic affinities between genets (Genetta) and the enigmatic genet-like taxa Osbornictis, Poiana and Prionodon (Carnivora, Viverridae). Zool. Scr., 33 (2), 117-129
5. Yoder AD, Burns MM, Zehr S, Delefosse T, Veron G, Goodman SM, Flynn JJ (2003) Single origin of Malagasy Carnivora from an African ancestor. Nature, 421 (6924): 734-737
6. Koepfli KP, Jenks SM, Eizirik E, Zahirpour T, Valkenburgh BV, Wayne RK (2006) Molecular systematics of the hyaenidae: relationships of a relictual lineage resolved by a molecular supermatrix. Molecular Phylogenetics and Evolution, 38,603-620

Chapter 3:

Nuclear SNP and ancient DNA refine the evolutionary history of the American Plains and Wood bison (*Bison bison bison*, *Bison bison athabascae*)

Nuclear SNP and ancient DNA refine the evolutionary history of the American Plains and Wood bison (*Bison bison bison*, *Bison bison athabascae*)

Julien Soubrier (Candidate)
Co-developed the research concept, performed the phylogenetic analyses, interpreted the results, developed the figures and wrote the manuscript.
I hereby certify that the statement of contribution is accurate.

Signed …… …… *Date* 02.04.12

Kefei Chen
Co-developed the research concept, processed the samples, performed ancient DNA research and library construction, co-produced the BovineSNP50 BeadChip genotypes, co-analysed the data, co-interpreted the results and co-wrote the manuscript.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… …… *Date* ……17/04/2012……

Michael S. Y. Lee
Co-developed the research concept, supervised the phylogenetic analyses and manuscript preparation.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… …… *Date* 17.4.12

Oliver Wooley
Co-developed the consensus building algorithm and helped analyse the data.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper ~~in the thesis~~.

Signed …… … *Date* 2.4.12

Simon Y.W. Ho
Co-interpreted the results, supervised phylogenetic analyses and manuscript preparation.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… …… *Date* ……05.04.2012……

Beth Shapiro
Provided samples, co-interpreted the results and supervised phylogenetic analyses.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… …… *Date* ……*3 Apr 2012*……

Wolfgang Haak
Supervised manuscript writing.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …. …… *Date* 02/04/2012

Steve Richard
Provided technical expertise on library preparation and data interpretation.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .. *Date* 02/04/2012

Robert Schnabel
Performed technical processing and analysis of SNP genotype libraries and co-interpreted the data.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… …… *Date* ……*April, 10, 2012*……

Jerry Taylor
Assisted with analysis of SNP genotype libraries and co-interpreted the data.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… …… *Date* …… *April, 10, 2012*…

Alan Cooper
Co-developed the research project, co-interpreted the results, co-wrote the manuscript and supervised the project.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… *Date* ……*05.04.2012*……

Nuclear SNP and ancient DNA refine the evolutionary history of the American Plains and Wood bison (*Bison bison bison*, *Bison bison athabascae*)

Julien Soubrier[1], Kefei Chen[1], Michael S. Y. Lee [2], Oliver Wooley[1], Simon Y.W. Ho[3], Beth Shapiro[2], Wolfgang Haak[1], Steve Richard[1], Robert Schnabel[6], Jerry Taylor[6] and Alan Cooper[1]

[1] Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, University of Adelaide, Adelaide, SA 5005, Australia;

[2] South Australian Museum, North Terrace, Adelaide, South Australia 5000, Australia

[3] School of Biological Sciences, University of Sydney, Sydney, NSW, Australia;

[4] Department of Biology, The Pennsylvania State University, University Park, Pennsylvania, United States of America;

[5] Division of Animal Sciences, University of Missouri, Columbia, MO 65211, USA.

ABSTRACT

Microarrays designed to simultaneously genotype large numbers of SNPs (Single Nucleotide polymorphisms) are increasingly available for commercial species. In addition to standard genetic screening, the rapid and cost-efficient access to large numbers of variable genetic markers distributed across nuclear genomes is of potential utility for evolutionary studies. However, microarray SNP genotyping data differs in the coding of alleles and the selection of variable markers within a species from traditional sequence data used in molecular phylogenetics, raising questions about the applicability of classical phylogenetic methods.

In addition, commercially available nuclear SNP microarrays also hold considerable potential to provide genome-wide data from ancient samples, and allow access to multiple individuals. However, the ability to generate reliable nuclear SNP genotypes from ancient samples is yet to be explored in detail.

Here, we attempt to generate nuclear SNP genotypes from 5 ancient specimens of Steppe bison (*Bison priscus*, extinct) using the Illumina BovineSNP50 assay, and to characterize potential biases due to degraded DNA. In addition, we used existing nuclear SNP genotypes from American bison, cattle and related bovid species to study the capacity of new phylogenetic methods to correctly infer past evolutionary history from microarray ascertained genotypes. On contrast to current conservation policies, we find a significant genetic divergence between Plains and Wood populations of American bison (*Bison bison bison* and *Bison bison athabascae* respectively), and propose a new evolutionary scenario for their origin.

INTRODUCTION

Importance of nuclear SNPs libraries

Genomic profiling via single nucleotide polymorphism microarrays (SNP chips) provides rapid and inexpensive access to genome-wide molecular data from large numbers of individuals. In addition to standard applications for medical or genomic screening, or commercial breed development within a single species, SNP chip genotype assays have recently been shown to have more widespread utility for evolutionary biology. SNP chips designed for one specific breed or species have been used to successfully screen the genomes of other taxa (Vonholdt et al. 2010 with dogs; Elferink et al. 2012 with chickens; McCue et al. 2012 with horses). At one extreme, a recent study of Bovidae demonstrated that SNP chips have the potential to construct large-scale phylogenies including different families (Decker et al. 2009).

Furthermore, SNP chip approaches have great potential in ancient DNA research where standard genomic sequencing approaches are complicated by the large amounts of exogenous DNA, mostly microbial, found in most ancient samples (Noonan et al. 2005). By targeting specific informative genomic loci and avoiding microbial DNA, SNP chips could provide access to endogenous genome-wide data within ancient populations at a reasonable cost and timeframe. This potential was recently demonstrated in the Bovidae study when a nuclear genotype library was generated for a 20,000 year old specimen of an extinct bison species (Decker et al. 2009).

However, to date, relatively few ancient DNA studies have applied SNP chips, and the growing field of paleogenomics has been based largely on the application of next generation sequencing approaches to whole mitochondrial and partial nuclear ancient genomes (Miller et al. 2008; Briggs et al. 2009; Green et al. 2010; Krause et al. 2010; Rasmussen et al. 2010). Similarly, there has been little development of methods to apply classical phylogenetic methods to SNP chip data from ancient or modern samples. Several logistical constraints in the SNP genotyping method have limited the adoption of SNP chip methods for phylogenetic analysis, with a key issue being that most currently available SNP chips use only two colour states to differentiate biallelic sites for reasons of cost effectiveness. In this system a single colour is used to represent two different nucleotides (e.g. A = a or t; B = c or g). This results in three possible character states for biallelic data: colour 1 for homozygote ‘AA’, colour 2 for homozygote ‘BB’ and a mix (third state) for heterozygote signal ‘AB’. The resulting three-state data would benefit from the use of an explicit substitution model, as implemented in maximum likelihood (ML) or related Bayesian algorithms. However, the large amount of data generated by SNP genome surveys (typically >50,000 SNPs) means that the limits of computational capabilities are rapidly reached for these methods. For this reason,

phylogenetic methods without complex substitution models such as simple parsimony or distance are more likely to provide results in reasonable time frames.

In addition to the methodological challenges imposed by the SNP coding system for phylogenetic studies of living or ancient taxa, the capability of commercially available SNP genotyping platforms to accurately recover sequence information from ancient genome amplifications is yet to be fully explored. In the present study, the initial results of Decker et al. (2009) were extended using additional samples of ancient Steppe bison (*Bison priscus*), and the results were used to assess the capacity of existing bovid genomic SNP chips to accurately genotype ancient individuals.

Background on bison evolution

Morphological taxonomy

Paleontological and morphological records have revealed a vast diversity of extinct bison species and sub-species, many proposed on relatively insecure morphological grounds. The paleontological record indicates the genus *Bison* has its origins in southern Asia (McDonald 1980). In the Late Pleistocene, the Steppe bison (*Bison priscus*) spanned the Holarctic from the UK to northern Mexico, while separate species were found in the New and Old Worlds. In North America, the long-horned bison (*Bison latifrons*) was succeeded by the ancient bison (*Bison antiquus*) while in Europe the Caucasus bison (*Bison bonasus caucasicus*) and the Carpathian bison (*Bison bonasus hungarorum*) are recognised (Guthrie 1970; Prusak, Grzybowski, Zieba 2004). Currently, only two species of bison are recognised: the European bison (*Bison bonasus*) and the American bison (*Bison bison*), which includes the Plains bison (*B. b. bison*) and Woods bison (*B. b. athabascae*) of Canada. The modern species are thought to be direct descendants of Late Pleistocene *B. antiquus* in America and *B. bonasus* in Europe, respectively.

American and European bison are morphologically similar (van Zyll de Jong 1986; Guthrie 1990) and are capable of hybridising. These characteristics have led to their current classification as closely related sister species (Skinner, Kaisen 1947; McDonald 1980).

Genetic markers

Mitochondrial DNA sequences from extant and extinct bison lineages have shown that the American bison is descended from eastern populations of the Steppe bison that had crossed the Bering land bridge to North America during the Late Pleistocene *ca.* >200 ky (Shapiro et al. 2004). Modern populations of both American and European bison show low mitochondrial genetic diversity, and this is thought to relate to known pronounced population bottlenecks in their recent history. The American bison was restricted to less than a hundred

individuals by the late 19th century (Roe 1970), while all living European bison descend from 17 animals originating from two small late-19th/early-20th century populations (Slatis 1960). Extant populations of European bison now exceed more than 2000 individuals, of which all pure-breed European bison represent the recombination of only 12 diploid sets of genes (Slatis 1960). Interbreeding with modern domestic cattle since the population bottlenecks has complicated the conservation of both American and European bison.

Nuclear and mitochondrial genetic analyses have reconstructed markedly different evolutionary relationships between European and American bison (Verkaar et al. 2004). Autosomal nuclear DNA studies confirmed the close morphological relationship between European and American bison (Buntjer et al. 2002; Vasil'ev et al. 2002; Verkaar et al. 2004), with both bison species forming a monophyletic clade closely related to Yak. On the other hand, mitochondrial phylogenies suggest the closest relative of the American bison is the Yak (*Bos grunniens*), while the European bison falls with cattle (*Bos taurus*)/Zebu (*Bos indicus*) (Janecek et al. 1996; Verkaar et al. 2004). Two hypotheses have been proposed to explain the difference between the phylogenies: incomplete lineage sorting, in which two distinct mitochondrial lineages survived in the bison/yak populations until the recent species-level split with European bison, or genetic introgression, where ancestral European bison mated with cattle/zebu individuals, perhaps during or shortly after a population bottleneck or with a male sex-bias so that ox/zebu-like mitochondrial DNA was able to become fixed within the population (Janecek et al. 1996; Verkaar et al. 2004). Neither explanation seems parsimonious, and the issue remains unresolved.

Plains and Wood American bison sub-species

The distinction between the living American Plains and Wood bison is also confused. The taxonomic distinction is based on phenotypic characters such as body size and coat characteristics (McDonald 1980; van Zyll de Jong et al. 1995), but this has been challenged and mitochondrial genetic data shows no separation between the two groups (Shapiro et al. 2004; Douglas et al. 2011; see also Halbert 2005 for a review). However, analysis of bison genomic data produced using the Illumina BovineSNP50 chip array, which analyses 54,693 variable cattle SNPs, generated a F statistic suggesting that there is indeed significant genetic differentiation between Plains and Wood bison (Pertoldi et al. 2010). This issue has significant conservation and agricultural implications in North America, and has become an important political consideration.

The BovineSNP50 chip

The BovineSNP50 chip was designed to genotype genomic SNPs known to be variable in different cattle breeds in a rapid, cost-efficient way (Matukumalli et al. 2009). A recent study (Decker et al. 2009) showed that the BovineSNP50 chip could be used to genotype animals across all five families of higher ruminants (Antilocapridae, Giraffidae, Moschidae, Cervidae, and Bovidae), and that a maximum parsimony analysis could produce a highly-resolved phylogeny that was consistent with accepted taxonomic groupings. Both Plains and Wood American bison were included in the study, but did not appear reciprocally monophyletic in the analysis and were consequently represented as one clade (American bison). In addition, one ancient specimen (a Steppe bison, *Bison priscus*) was genotyped, and appeared basal to the American bison in the phylogeny. However, maximum parsimony does not incorporate explicit models of sequence evolution, and is unlikely to adequately account for complexities in the SNP data (e.g. unequal base frequencies and substitution rate heterogeneity across sites). A key issue is that heterozygote alleles (AB) are not easily interpreted within parsimony analyses, and were thus removed from the analysis.

Ascertainment bias

The BovineSNP50 targets SNPs known to be variable between cattle breeds (*Bos taurus*), but relatively few of these are variable in other species and, as a result, far fewer differences are observed between individuals. In fact, the proportion of polymorphic sites and the level of heterozygosity decrease with increasing genetic distance between cattle and the species of interest, so that genetic distances demonstrate an ascertainment bias. For example, Zebu (*Bos indicus*) has a lower proportion of polymorphic sites than any *Bos taurus* breed and would consequently have a shorter branch on an inferred phylogeny, irrespective of the true evolutionary distances and branch lengths.

The ascertainment bias of the chip has two main consequences for phylogenetic studies:

1. The non-proportionality of estimated genetic distances between taxa.
The ascertainment bias might not necessarily impact the topology (branching order) of inferred phylogenetic trees, but would definitely bias the estimated branch lengths with increasing genetic distance from cattle. However, within a single clade where all taxa are equidistant to cattle, it is likely that the relative genetic distances between taxa would be proportional to evolutionary time. For example, the Gaur (*Bos gaurus*), Banteng (*Bos javanicus*), American bison (*Bison bison*), European bison (*Bison bonasus*) and Yak (*Bos*

grunniens) form a monophyletic group within Bovidae that is sister to cattle and Zebu. In that regard, SNP data from the former five species should share a similar level of ascertainment bias, such that genetic distances could be compared between them. As comparable genetic distances are a requirement for molecular dating, the ascertainment bias means that dated phylogenies can only be calculated for monophyletic clades that are outgroups to cattle breeds.

2. The number of polymorphic sites is lower for taxa that are evolutionary distant from cattle, limiting phylogenetic signals between closely related individuals. This problem is further compounded if heterozygote alleles are ignored by the phylogenetic method. In the case of American bison, the reduced amount of potentially informative characters complicates population level studies and comparisons of the Plains and Wood bison populations (Decker et al. 2009).

In the present study, we used maximum likelihood analyses to re-analyse the bovid SNP chip data from Decker et al. (2009) and utilise phylogenetic signal from heterozygote alleles in the SNP libraries. We re-calculate phylogenies for both the cattle and bison groups, and examine the taxonomic status of Plains and Wood bison as different species or sub-species.

In addition we genotyped five ancient Steppe bison specimens with the BovineSNP50 chip to explore the potential of commercially developed SNP arrays to recover evolutionary information from ancient DNA.

METHOD

Data

We used the data set of 40,843 SNPs selected by (Decker et al. 2009) based on their autosomal location, high call rate and polymorphic state between cattle breeds. This was a subset of the 54,693 SNPs carried on the microarray.

Modern samples

In addition to the genotype libraries used by (Decker et al. 2009) for Cattle, Zebu, American bison, Yak, Gaur and Banteng, an additional seven libraries of European bison (*Bison bonasus*), from the Wisentgehege Springe in Germany, were included (Personal communication).

Ancient samples

In addition to the single Siberian Steppe bison (BS662, three replicates), four additional Steppe bison were genotyped (see supplementary table 1): a 20 ky Canadian sample from Yukon (A3133, nine replicates), two 30 ky Chinese samples from Jilin (A5880 and 5881, one replicate each) and a 50 ky Russian sample from the Urals (A002, two replicates).

Ancient DNA extraction and amplification

Ancient DNA was extracted from fossil bison bone specimens using the standard phenol/chloroform/Amicon Ultra-4 method (17) and silica powder matrix extraction protocol (Austin et al., in preparation). DNA extractions, omniplex library preparations, and PCRs were set-up and performed in a geographically isolated, dedicated ancient DNA facility at the University of Adelaide, Australia. Sixteen libraries were selected from a panel of 60 whole genome amplified libraries from ancient DNA were created using different whole genome amplification methods (see Suppl. Table 1). Whole genome amplified ancient DNA products were finally purified using the GenElute PCR Clean-Up kit (Sigma-Aldrich) or ethanol. Ancient DNA libraries were verified by PCR amplification and sequencing of the hypervariable mtDNA control region before analysis with the BovineSNP50 BeadChip (Illumina).

WGA2 (GenomePlex® Whole Genome Amplification Kit): To generate a library of genomic fragments from limited ancient DNA extracts, DNA was amplified using the PCR-based GenomePlex® Whole Genome Amplification kit (WGA2; Sigma-Aldrich) according to the following protocol: 10 μL DNA were thoroughly mixed with 2 μL library preparation buffer and 1 μL library stabilization solution, and denatured at 95 °C for 2 min. After denaturation, 1 μL library preparation enzyme was added to generate omniplex libraries, followed by a series of incubations at 16 °C for 20 min, 24 °C for 20 min, 37 °C for 20 min, and 75 °C for 5 min in a thermal cycler (Corbett Life Science). The omniplex libraries were amplified using a limited number of genomic amplification cycles. PCR amplification was conducted in a 75 μL reaction volume containing 14 μL omniplex library, 7.5 μL amplification master mix, 48.5 μL nuclease-free water, and 5 μL WGA DNA polymerase. The PCR amplification conditions were initial denaturation at 95 °C for 3 min, followed by 15 cycles of 94 °C for 15 s and 65 °C for 5 min.

WGA4 (GenomePlex® Single Cell Whole Genome Amplification Kit): DNA was amplified according to the manufacturer's protocol but with the removal of the single cell

lysis and fragmentation steps (Sigma-Aldrich). 2 µL of 1× single cell library buffer and 1 µL of library stabilization solution was added to 10 µL of ancient DNA. Following mixing the solution was placed in a thermal cycler for denaturation at 95 °C for 2 min, and cooled on ice. After denaturation, 1 µL of library preparation enzyme was added, and incubated as follows: 16 °C for 20 min, 24 °C for 20 min, 37 °C for 20 min, 75 °C for 5 min and 4 °C hold. The following reagents were added to the entire 14 µL reaction: 7.5 µL of 10 × amplification master mix, 48.5 µL of nuclease-free water, 5.0 µL of WGA DNA polymerase. The PCR amplification conditions were 95 °C for 3 min, followed by 20 cycles of 94 °C for 30 s, 65 °C for 5 min, and hold at 4 °C.

Repli-g® FFPE: This system used the kit for whole genome amplification of DNA from Formalin-Fixed Paraffin-Embedded (FFPE) tissue (QIAGEN). A 1 × FFPE master mix was prepared by adding 8 µL FFPE buffer, 1 µL ligation buffer, and 1 µL FFPE enzyme. 10 µL of master mix was added to 10 µL of DNA, mixed and centrifuged briefly. The sample was incubated at 24 °C for 30 min, 95 °C for 5 min, and cooled down to 4 °C in a thermal cycler. After adding 3 µL of Repli-g master mix to the denatured DNA, the solution was mixed and centrifuged briefly before incubation at 30 °C for 8 hr, 95 °C for 5 min and storage at 4 °C.

Character composition and consensus building

To survey differences in character composition between genotype libraries, three parameters were considered: the ratio of SNP homozygote calls (AA/BB); the proportion of missing data (?); and the SNP heterozygosity, after accounting for missing data (AB/(AA+AB+BB)). The character composition was observed to vary considerably within the ancient libraries, particularly the proportion of missing data.

Consensus sequences were constructed from the replicates of the ancient libraries in order to separate artifactual signal, randomly attributed between replicates, from genuine calls, consistent between replicates. The individual genotypes were aligned, and consensus sequences were built from six individual replicates of the Steppe bison sample A3133 (Yukon, Canada), selected based on their character composition. Different quality thresholds have been set to generate a consensus call for each individual site (column in the alignment). The first consensus sequence threshold assigned a missing data '?' call to sites where the proportion of missing data was equal to or greater than 75% of the replicates (i.e., '?' in more than four of the six replicates), and was termed 'Consensus 75'. A second, more conservative, consensus threshold was constructed using a cutoff of >50% ('Consensus 50'), where the site was assigned '?' unless a SNP was called in more than three of the six replicates. These

thresholds were used to perform an initial filter of the data, before consensus calls were made for the remaining sites. A consensus call was made for sites only where a single SNP allele (e.g. AA) or ? was observed between all the replicates. In all other cases the site was assigned as missing data, effectively discarding sites where any form of conflicting signal existed between replicates of the same sample. This conservative approach was adopted to clean the ancient data from potential noise generated by sequencing errors. To further investigate this issue both consensus libraries were analysed with, and without, heterozygote SNP alleles (AB) in the phylogenetic analyses.

Phylogenetic analysis

The program RaxML v7.2.8 (Stamatakis 2006) was used to perform maximum likelihood tree searches on multi-state data sets. The three characters from the Bovine50SNP chip (AA, BB and AB) were considered as different states in an explicit analogue of the General Time Reversible (GTR) substitution model, with separate substitution parameters for the three possible transformations (AA-BB; AA-AB; BB-AB). For all analyses, 20 maximum likelihood searches were conducted to find the best tree, and branch support was estimated with 500 bootstrap replicates using the rapid bootstrapping algorithm (Stamatakis, Hoover, Rougemont 2008).

Three data sets were assembled for phylogenetic analysis using data from (Decker et al. 2009) (1) Domestic cattle: Re-analysis of 371 cattle genotypes covering three Indian breeds (*Bos indicus*), one African, two Asian and 41 European breeds (*Bos taurus*).

(2) Bison and relatives: Analysis of the Bos/Bison clade with the additional European bison sequences, featuring 10 Gaur, four Banteng, two Yak, seven European Bison, 25 Plains bison, and 30 Wood bison individuals. Character composition analyses revealed several specimens with apparent low data quality (one Plains and eight Woods bison), which were removed from the analyses. In addition, individual replicates and consensus sequences of the five ancient Steppe bison samples were analysed.

(3) Bison: To examine the relationships within and between American and European bison, 24 Plains bison, 22 Wood bison and seven European bison individuals were analysed.

In addition, a dated phylogenetic analysis was performed on a data set comprising 54 Yak and American bison individuals to investigate the potential to create a temporal timescale for bovid evolution using SNP chip data. The program BEAST v1.6.2 (Drummond, Rambaut

2007) was used with the *BEAST option to account for both inter and intra-specific diversity. Based on previous results, Plains and Wood bison were considered as two reciprocally monophyletic populations for the analysis. A 3-state general substitution model was implemented to allow for different equilibrium state frequencies and different substitution rates between the three possible substitution types. A gamma distribution with 6 rate categories was used to account for rate heterogeneity among sites. The tree was calibrated by setting the divergence between Yak and Bison at around 2.5 million years (My), using a lognormal distribution with a mean of 2.5 My and 95% of the prior probability between 2 and 3 My (Wang et al. 2010). An uncorrelated lognormal relaxed-clock was used to account for rate heterogeneity among lineages (Drummond et al. 2006). Considering the large size of the data set, eight MCMC chains of the same analysis were run to check for convergence toward the same likelihood. Each chain was run for 100,000,000 iterations, with posterior samples drawn every 10,000 steps. The convergence of the analysis was checked in Tracer 1.5 and the initial 10% was discarded as burn-in (Rambaut, Drummond 2007).

Finally, to check on the placement of two of the Steppe bison samples genotyped for nuclear SNPs on the mitochondrial phylogenetic tree, a dated phylogenetic analysis was performed using 191 American and Steppe bison mitochondrial control regions. This data set is composed of 189 sequences form Shapiro et al. (Shapiro et al. 2004) and control region sequences from the both the Yukon Steppe bison A3133 and the Siberian Steppe bison BS662. The substitution model HKY + G6 + I was selected by comparison of Bayesian Information Criterion scores in ModelGenerator v0.85 (Keane et al. 2006). The phylogenetic calculation was conducted with the program BEAST 1.6.2 (Drummond, Rambaut 2007) using a coalescent approach. The GMRF Skyride model (Minin, Bloomquist, Suchard 2008) was used to account for variations in the population demographic history of bison, and a relaxed uncorrelated lognormal clock was used to account for potential rate variations between lineages. Calibrated radiocarbon dates associated with all sequences were used as calibration points. The Tip Date Sampling option (Shapiro et al. 2011) implemented in BEAST, was used to account for the age uncertainty of three samples (BS129, BS368 and BS662): a lognormal distribution prior was set with minimum bound of 0 and 95 % of the prior probability lower than 2 ky for BS129 and BS368, and a normal distribution prior was used for BS662 with a mean of 20 ky and a standard deviation of 10 ky. A ‘date randomization test’ has already been conducted on the same dataset (at the exception the two additional samples BS662 and A3133) to check that the signal from the radiocarbon dates associated with the ancient sequences was sufficient to calibrate the tree (Ho et al. 2011b).

RESULTS AND DISCUSSION

Using a probabilistic reconstruction method (ML) to recover phylogenetic signal from heterozygote characters

In the original analysis of the Bovine50SNP chip data, it was noted that maximum parsimony (MP) methods could only reconstruct the biogeographical history of cattle breeds once heterozygote SNP calls (AB) were removed from the analysis (Decker et al. 2009). Unlike maximum likelihood or Bayesian based methods, maximum parsimony does not incorporate explicit substitution models. The costs (and thus the frequencies) of all possible changes between states AA, BB, and AB are considered identical. In this situation, a substitution between homozygote characters (AA-BB) has the same cost as a substitution between homozygote and heterozygote characters (AA-AB and BB-AB), which clearly might not reflect the evolutionary path length.

The use of probabilistic methods such as ML or Bayesian approaches allows the implementation of explicit models of evolution that account for different substitution probabilities between the three character states of the SNP libraries. However, most programs used to perform ML or Bayesian calculations (like PhyML or MrBAYES) could not complete phylogenetic inferences on such large data sets (~50,000 characters for up to 600 taxa) in a reasonable amount of time. The Randomized Accelerated Maximum Likelihood (RaxML) program has been developed for this situation, and is built around an optimized version of the rapid hill-climbing algorithm (Stamatakis 2006). Combined with newly developed rapid bootstrap heuristics, and using parallelized computing, this method has calculated some of the largest ML phylogenies to date (Stamatakis, Hoover, Rougemont 2008). In the present study, the program RaxML has been used with a GTR-like substitution model specifically developed for multi-state data sets. This method calculates an explicit substitution model empirically from the data, allowing for unequal equilibrium character state frequencies and rates between different types of substitutions.

When the phylogeny of 48 cattle breeds was re-calculated including the signals from heterozygote characters via the parameters described above, most of the biogeographical history described in the original analysis was recovered, with equal or higher support for the main clades (Figure 1). However, several key differences are apparent:

- Only three breeds appear paraphyletic (Salers, Corriente and Angus), compared to seven in the original MP tree (Salers, Corriente, Angus, Hanwoo, Texas Longhorn, Limousin, Maine-Anjou).
- The three Italian breeds fall basal to other European cattle, rather than the New World Spanish breeds observed in the original MP tree. This arrangement is a much better match for the movement of domesticated cattle from the Fertile Crescent along the

Mediterranean Coast and into SW Europe during the Neolithic, and removes a phylogeographic inconsistency identified in the original analysis.

- All British cattle breeds form a monophyletic group, with Irish representatives at the base of the group.

These findings improve the match between phylogeny and biogeographic history, and confirm that heterozygote sites can contribute phylogenetic signals when an adequate substitution model is used.

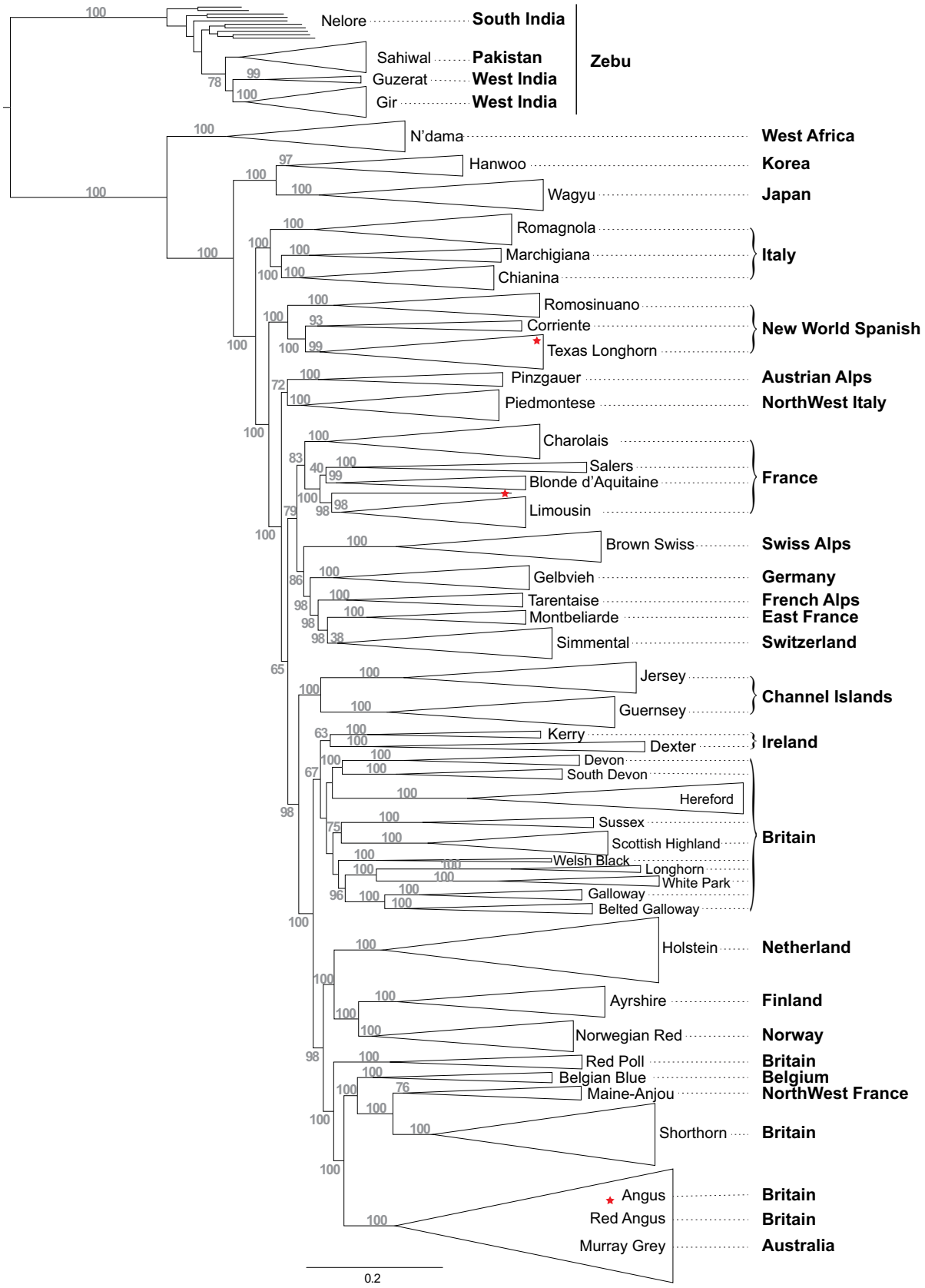


Figure 1. Phylogenetic tree of 48 Cattle breeds and their geographical origin. Four Zebu lineages were used as an outgroup, and heterozygotes were included in the calculation. Grey numbers represent the bootstrap support values and red stars the paraphyletic groups.

Character composition and consensus building

Modern bison composition and AB bias

Most modern bison have a near identical homozygote ratio (AA/BB = ~0.63), and low proportions of missing data (~ 2 to 4%) and heterozygosity (AB = ~ 0.8 to 2%) (Figure 2). However, eight Wood and one Plains bison SNP chip libraries contain an abnormally high proportion of missing data for modern samples (5 to 36%). As this suggests potential problems with the quality of the DNA, the nine libraries were removed from all analyses. Interestingly, the elevated proportion of missing data in these samples was associated with a homozygote ratio deviating greatly from the value observed for other modern bison samples (0.63), and a high proportion of heterozygotes. The association between low quality samples and high proportions of heterozygotes may be due to the normalization of colour signals performed during the processing of the SNP chip data. If the signals are noisy due to poor DNA hybridization between probe and target sequence, the normalization of colour intensity is more likely to artifactually call a heterozygote (bi-colour) than a homozygote (monochrome).

Plains/Wood heterozygosity

When compared to Plains bison, Wood bison bison SNP chip libraries show reduced levels of heterozygosity. Given the recent history of population bottlenecks in both groups, this might suggest that Wood bison suffered a more dramatic or prolonged period of restricted population size. Alternatively, some Plains bison are known to have levels of introgressed cattle DNA which would increase heterozygosity. However, the samples analysed here are thought to contain no detectable levels of introgressed cattle DNA (Decker et al. 2009).

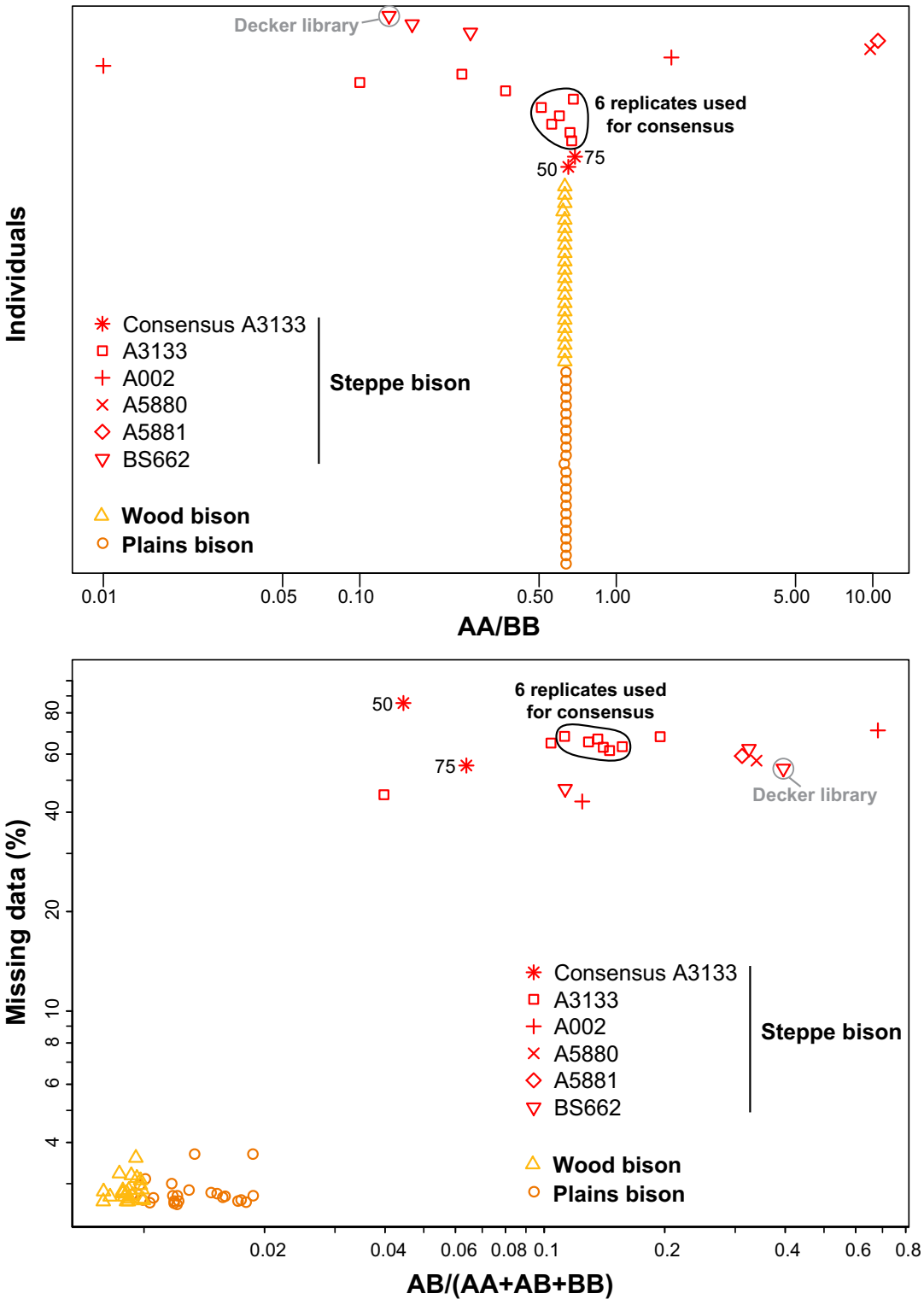


Figure 2. Character composition plots. The top graph represents the ratio between the two homozygote characters for each library of American bison (orange, each replicated library of ancient bison samples (red) and the consensus sequences (red star). The second graph shows the relationship between the proportion of missing data and the proportion of heterozygotes for the same genotype libraries.

Ancient libraries

The SNP chip libraries generated from the ancient bison samples contain high proportions of missing data (43 to 68%), as expected. As noted in the poor quality modern libraries, the elevated proportions of missing data are strongly associated with high levels of heterozygosity (AB). Most ancient libraries showed a strong bias away from the AA/BB ratio of 0.63 observed in modern bison, and ranged from 0.01 (99% of homozygote calls are BB) to 10.48 (90% of homozygote calls are AA). Only six library replicates from the Yukon Steppe bison sample A3133 had homozygote ratios close to modern bison, and ranged from 0.51<AA/BB<0.68. As a consequence, these six libraries were used to build the consensus sequence for the Yukon Steppe bison.

Consensus building

Due to the apparent low quality of several of the ancient bison libraries, only sites without conflicting calls between replicates were considered robust and used in the consensus sequence. This conservative approach was used to remove as much noise as possible from the genotype libraries, and is feasible due to the high number of characters available on the Bovine50SNP chip. The different calling thresholds used for consensus building resulted in 22,621 positions being labelled as missing data (55.4%) for Consensus75, and 34,965 positions (85.6%) for the more conservative Consensus50. However, that still leaves 18,222 and 5,878 positions, respectively, for phylogenetic analysis.

Impact of consensus building on character composition

Interestingly, both consensus libraries converged towards the AA/BB ratio observed in modern genotypes (~0.63 for modern bison, Figure 2), with Consensus75 at 0.69 and Consensus50 at 0.65. Furthermore, the consensus building approach drastically reduced the amount of heterozygote calls relative to individual ancient library replicates, although the values for both consensus libraries remained much higher than those observed for modern samples (Figure 2). To assess the impact of this, the phylogenetic analyses were performed with and without all heterozygote calls removed only from the two ancient consensus sequences (i.e., heterozygote calls in modern taxa were retained, due to their informativeness discussed above).

Bison phylogeny

Position of European bison

The European bison (*Bison bonasus*) was found to be closely related to American bison (with 100% bootstrap support, Figure 3a) in the maximum likelihood phylogeny, as

previously observed with nuclear sequences (Buntjer et al. 2002; Verkaar et al. 2004). This is consistent with morphological and phenotypic analyses, but directly contrasts with mitochondrial studies (Janecek et al. 1996; Verkaar et al. 2004). However, because the BovineSNP50 chip data is drawn from across the entire nuclear genome it surveys the species history far more than the mitochondrial DNA, which represents a single non-recombining locus. The SNP data does not provide evidence about why the evolutionary history of the mitochondrial DNA deviates from that the likely true species tree, and further ancient DNA studies of the European bison will be needed to investigate this issue.

Genetic differentiation of Plains and Wood American bison

Surprisingly, within the diversity of American bison, the sampled Plains and Wood bison individuals clearly divided into two reciprocally monophyletic clades (Figure 3). Within a phylogenetic tree comprising only bison species (Figure 3b), a deep genetic split was observed supported by strong bootstrap support (100% for Wood bison and 93% for Plains bison). The reciprocal monophyly was also strongly supported when the nine low quality modern bison samples were included in the analysis (Supplementary Figure 1). As noted in the 2009 analysis, when heterozygote characters were excluded (Supplementary table S1 of Decker et al. 2009) reciprocal monophyly was not detected (Supplementary Figure 2). However, this was because just one Plains bison individual fell basal to both groups, although this position had low support values. This appears to suggest a relatively deep genetic separation between these two groups, and low levels of introgression at least in the sampled Woods bison populations. This is perhaps surprising given the extensive co-habitation between the two groups in some national parks, and suggests that further conservation management is required. While it is difficult to relate SNP genetic diversity to taxonomic status, the size of the genetic separation and reciprocal monophyly would at least merit sub-species status, and is compatible with a species-level split.

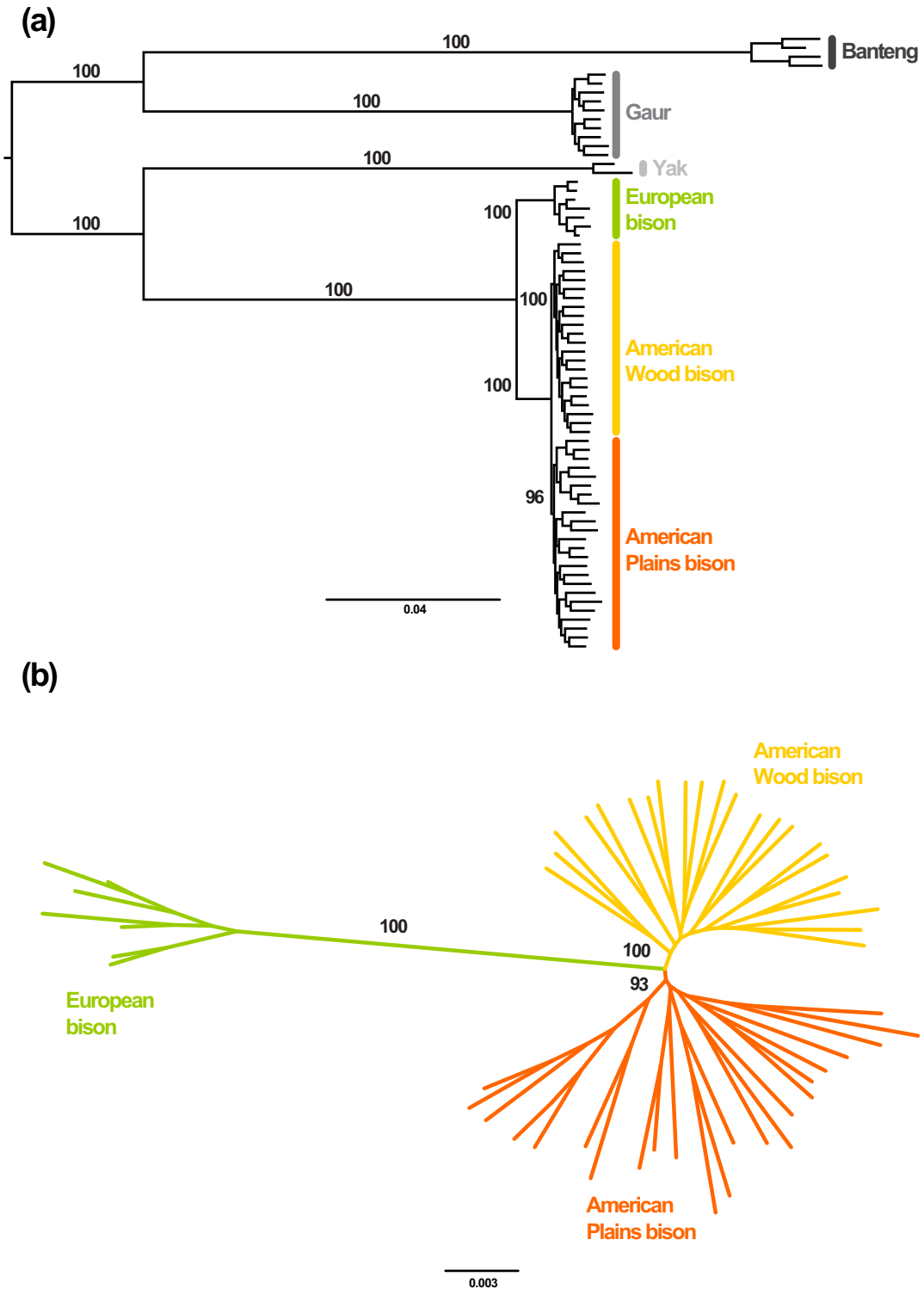


Figure 3. Phylogenetic placement of European and American bison based on 40,843 nuclear SNPs. The Multi-state option of RaxML was used to account for all three possible SNP characters, including heterozygotes. Black numbers represent branch support (based on 500 bootstrap replicates). **(a).** Phylogeny including three outgroup species: the Yak (*Bos grunniens*), the Gaur (*Bos gaurus*) and the Banteng (*Bos javanicus*). European bison were retrieved as sister taxa to both American bison clades, and there is strong support for reciprocal monophyly between American Plains and Wood bison. **(b).** Phylogenetic tree including only European and American bison, showing a clear genetic divergence between Plains and Wood bison clades.

Position of Steppe bison

When the consensus sequences from the six high-quality replicates of the ancient Yukon Steppe bison sample A3133 were included in the phylogeny, the analysis systematically placed the ancient specimen within the diversity of modern Wood American bison (Figure 4). Phylogenetic trees were constructed with the two different consensus sequences created for the Yukon Steppe bison, both heterozygote characters in the Steppe bison either included or excluded. With the Consensus75 sequence, the Steppe bison was placed within the Wood clade with very high bootstrap values (98-99 %, Figure 4a). A similar topological position was retrieved with the much more conservative Consensus50 sequence, but with a shorter branch length and lower bootstrap support (62-66 %, Figure 4b). The shorter branch length suggests that the stringent parameters used to build the Consensus50 are indeed removing noise, but the lower bootstrap signal indicates that some phylogenetic signal is lost as well. For both consensus libraries the removal of heterozygote characters drastically reduced the branch length of the ancient sample, suggesting the atypically high level of heterozygote calls obtained from ancient extracts is likely to be artefactual, as expected from the composition plots.

When all nine library replicates of the Yukon Steppe bison were used individually (without heterozygotes) to calculate the its phylogenetic position, seven fell in a similar position within the Wood bison clade (Supplementary Figures 3a/b/c). One replicate fell basal to all American bison, and another within the Plains bison diversity, but both topologies were associated with low support (<60 %) for both the American and Wood bison clades. Similar results are obtained when heterozygote characters are included, with longer branch length. Also, the individual replicates had notably long branch lengths relative to the consensus libraries, and together with the variable positions for the two outliers, confirms the utility of a consensus building approach when dealing with SNP chip analyses of ancient samples.

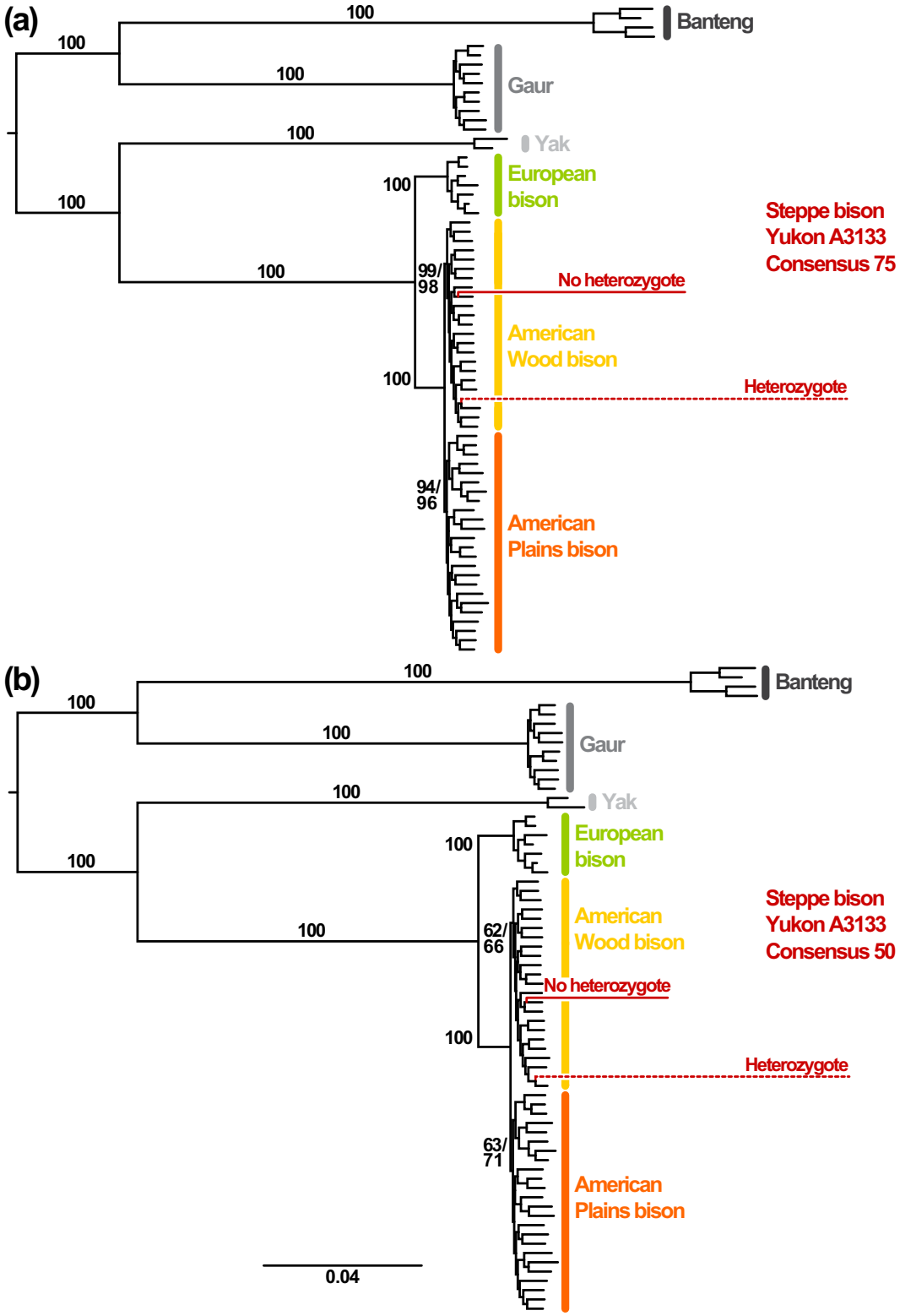


Figure 4. Phylogenetic placement of the Steppe bison (*Bison priscus*) A3133 from the Yukon, based on 40,843 nuclear SNPs and using different consensus sequences, with and without heterozygote positions. Dotted red branches indicate the position of the Steppe bison sample when heterozygote characters are included, while plain red branches indicate the position when heterozygotes are removed from the consensus. For both consensus sequences, heterozygote characters from the ancient samples are shown to produce longer branch length. **(a).** Using 'Consensus 75' (with a missing data threshold of 75% between replicates), there is strong support for the American Wood bison including Steppe bison, but a long branch for the ancient sample indicates an important amount of noise within the ancient libraries. **(b).** Using the more conservative 'Consensus 50' sequence, the branch lengths show less noise, but there is also a reduction in the phylogenetic signal with lower branch support for both clades of American bison.

The three library replicates of the Siberian Steppe bison BS662 show consistent placement at the root of American bison (both Wood and Plains) when analysed without heterozygotes, and inconsistent topological position when heterozygotes are included (Supplementary Figure 4). Although the character composition of all three BS662 SNP libraries suggests that the genotyping is less reliable than for the sample A3133, the inferred position of BS662, basal to all American bison, is consistent with the result published in 2009 (Decker et al. 2009).

For the three other Steppe bison samples, the topological position cannot be accurately determined with the SNP libraries (Supplementary Figure 5 and 6), as the Steppe individuals are grouping with the outgroup clades (Gaur and Banteng) or have low support values for the clade they fall in.

Dated tree

To generate a timescale for bison evolution, and estimate the timing of the genetic separation between Plains and Wood American bison, a dated phylogenetic analysis was performed using Yak, and American and European bison genotype libraries. Although the position of the Yukon Steppe bison within Wood bison has strong statistical support, the artificially long branch length created by the elevated heterozygote levels is likely to confound dating attempts and consequently the ancient sample was removed from the dated analysis. To circumvent the issue of ascertainment bias inherent in the SNP genotyping, dating was performed only on taxa within a monophyletic clade were all taxa were equally phylogenetically distant from cattle.

Four of eight runs converged toward the same mean likelihood, and the results were consequently combined to build the consensus tree (Figure 5). The combined statistics show convergence of the MCMC chains for all parameters, with effective sample size (ESS) values above 200. The other chains appeared to get stuck in a local maximum likelihood, and would potentially require more than 100,000,000 iterations to reach admixture. The time of the most recent common ancestor (tMRCA) of all bison (European and American) was calculated to be 594 ky BP (279 – 984), with a date of 308 ky BP (158 – 491) for the split between Wood and Plains bison.

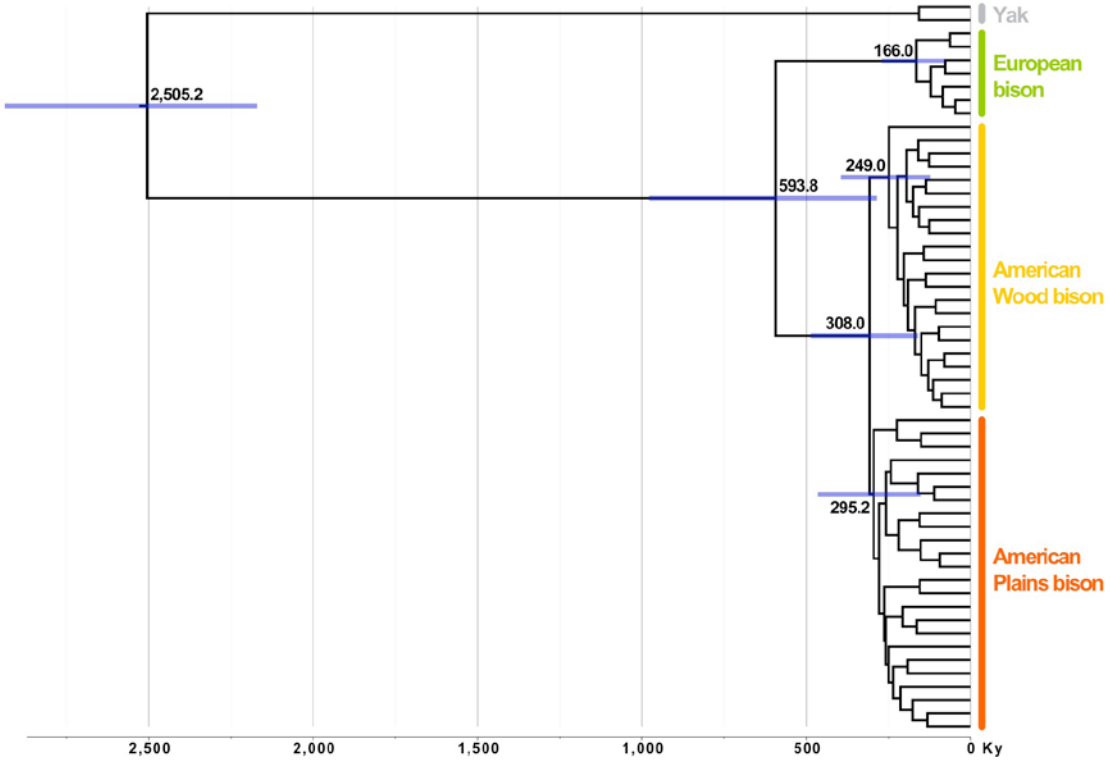


Figure 5. The maximum clade credibility tree estimated using Bayesian analysis of 55 bison and Yak nuclear SNP libraries. Estimated divergence times of the main clades are noted at the node and the blue bars represent confidence intervals (95% HPD).

The different topological signal obtained from mitochondrial and nuclear data makes it difficult to compare the estimates here with previous values based on mitochondrial sequences (Shapiro et al. 2004; Ho et al. 2008a). However, it is important to consider that the nuclear SNP dates are more likely to be impacted by the temporal dependency of molecular rates (Ho et al. 2005; Ho et al. 2011a), and are likely to be overestimates. This is because it is not possible to use the radiocarbon dates associated with the ancient samples to calibrate the SNP chip phylogeny, due to the potential impact of their artificially long branches. As a result, the temporal calibration is based only on a deep fossil calibration point estimated at 2.5 My BP, although it is known that the difference between calibration point and date of interest has a major impact on the accuracy of estimated dates (Ho et al. 2008b). In that regard, the tMRCA presented here should be considered as upper bounds. Although the present calculation demonstrates the possibility of using nuclear SNP data for molecular dating, additional calibration points will be required to further resolve the timescale of bison evolution.

Evolutionary history of American bison

Based on the paleontological record, Eurasian bison populations are thought to have reached the American continent through Beringia between 300 and 130 ky BP (McDonald 1980). Ancient and modern mitochondrial DNA sequences revealed that the modern

American bison originated from the physical separation of Steppe bison populations by ice sheets forming over North America between 25 and 13 ky BP (Shapiro et al. 2004). The entire mitochondrial genetic diversity observed in modern American bison (Wood and Plains) comes from the population restricted to areas south of the ice corridor during the last glacial maximum (LGM, clades 1a and 1b on Figure 6). This scenario is in accordance with the placement of the Siberian Steppe bison individual (BS662) basal to all existing American bison, however, it is in direct contrast to the deep genetic split observed between Plains and Wood nuclear genomes, and the position of a Steppe bison within the Wood bison diversity. Indeed, the latter suggests that Wood bison may actually be a surviving form of the Steppe bison populations north of the ice sheets during the LGM. This scenario is supported by the retrieval of a Steppe bison mitochondrial haplotype in a recent bison sample in Canada (50 to 300 C14 years BP, see supplementary data in Shapiro et al. 2004).

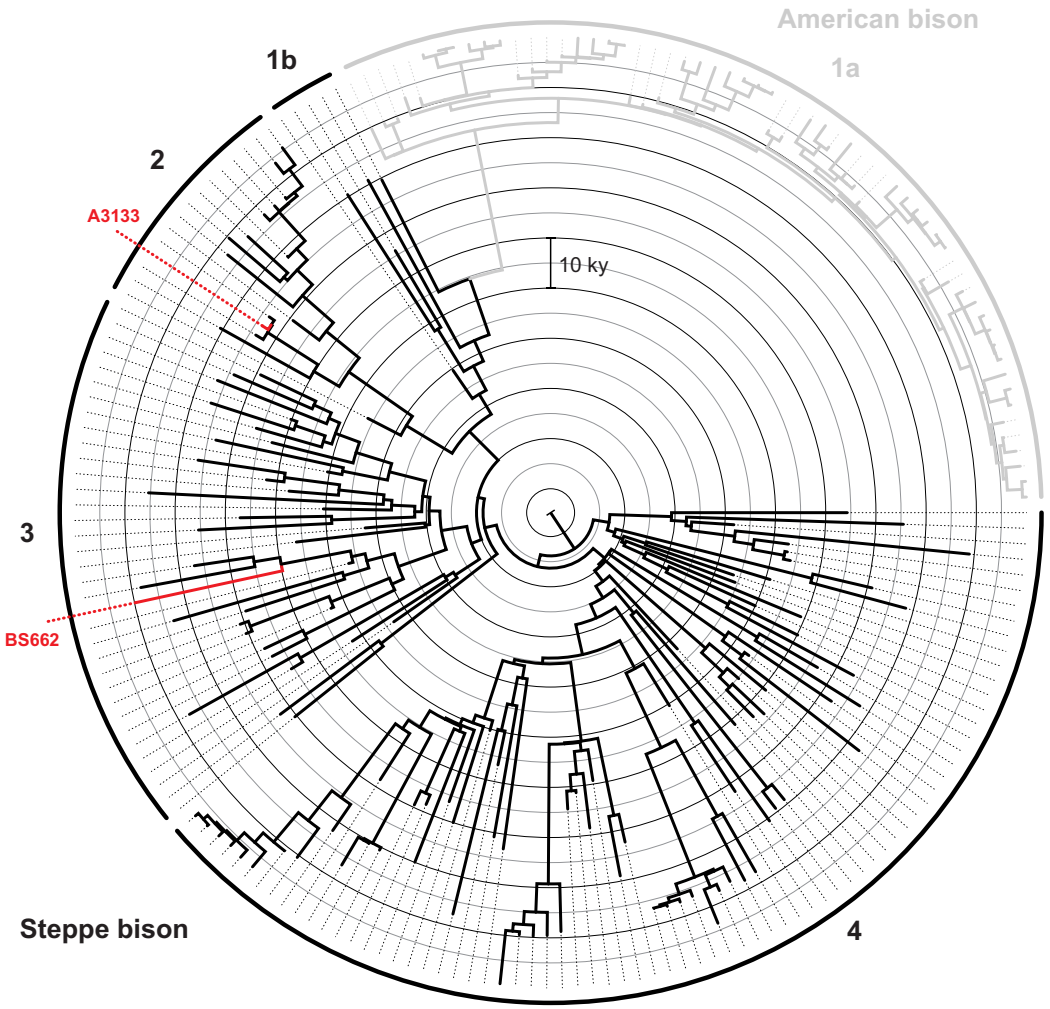


Figure 6. Maximum clade credibility tree showing the placement of both Steppe bison specimens (A3133 from the Yukon and BS662 from Siberia) on the mitochondrial phylogeny of Steppe and American bison. This dated phylogeny was calculated in BEAST v1.6.2 using the control region data set from Shapiro et al. (2004). Radiocarbon dates associated with the samples were used to calibrate the tree. Numbers correspond to the clade description from the original publication.

Furthermore, the Yukon Steppe bison (A3133) that falls within the Wood bison diversity on the nuclear phylogeny belongs to clade 2 on the mitochondrial phylogeny (Figure 6), immediate sister group to clade 1 containing all modern American bison mitochondrial haplotypes. On the other hand, the Steppe individual BS662, falling basal to all modern bison in the nuclear phylogeny, belongs to the mitochondrial clade 3, which has a more basal position below clade 2 and the modern haplotypes. One explanation to reconcile nuclear and mitochondrial phylogenetic signals involves recent mtDNA introgression. Under this scenario, the split between mitochondrial clades 2 and 1 (around 73 ky based on tip calibration) corresponds to the split between Plains and Wood bison on the nuclear phylogeny (estimated at 158 to 491 ky based on fossil calibration). Then individuals from the mitochondrial clade 2, including A3133, would be ancestors of the Wood bison 'sub-species'. The putative position of BS662 outside both Plains and Wood nuclear diversity would suggest that the mitochondrial clade 3 did not participate to the actual diversity of Wood bison. This is supported by the fact that most the individuals within clade 3 were shown to be part of the re-colonization wave toward China and Siberia from Alaska (Shapiro et al. 2004). This early separation between Plains and Wood bison populations, supported by the nuclear data, strongly differs from the common assumption, based on mitochondrial genetic diversity, that the Wood type of American bison is a recent phenotypical adaptation of the original Plains bison type, in response to the environment. If the scenario presented here is correct, then the absence of mitochondrial differentiation between Plains and Wood individuals could be explained by hybridization event(s), with introgression of Plains bison individuals (from south of the ice) into the Wood population (from north of the ice) through the ice-free corridor (after 14 ky BP, see Figure 1.D in Shapiro et al. 2004). The apparent fixation of the 'southern' mitochondrial type in all modern American bison may have been facilitated by the strong population bottleneck observed hundred years ago, caused by excessive hunting.

Incomplete lineage sorting has also been suggested to explain topological discrepancies between nuclear and mitochondrial markers in the European bison. However, the extensive sampling of ancient mitochondrial diversity in North American bison through time makes it likely that incomplete sorting of mitochondrial lineages would have been detected. It will be necessary to apply SNP chip methods to further ancient bison specimens to reconstruct the history of Woods and Plains bison in North America.

References

Briggs, AW, JM Good, RE Green, et al. 2009. Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science 325:318-321.

Buntjer, JB, M Otsen, IJ Nijman, MTR Kuiper, JA Lenstra. 2002. Phylogeny of bovine species based on AFLP fingerprinting. Heredity 88:46-51.

Decker, JE, JC Pires, GC Conant, et al. 2009. Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. Proceedings of the National Academy of Sciences of the United States of America 106:18644-18649.

Douglas, KC, ND Halbert, C Kolenda, C Childers, DL Hunter, JN Derr. 2011. Complete mitochondrial DNA sequence analysis of Bison bison and bison-cattle hybrids: function and phylogeny. Mitochondrion 11:166-175.

Drummond, AJ, SYW Ho, MJ Phillips, A Rambaut. 2006. Relaxed phylogenetics and dating with confidence. Plos Biology 4:699-710.

Drummond, AJ, A Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC evolutionary biology 7:214.

Elferink, MG, HJ Megens, A Vereijken, X Hu, RP Crooijmans, MA Groenen. 2012. Signatures of selection in the genomes of commercial and non-commercial chicken breeds. Plos One 7:e32720.

Green, RE, J Krause, AW Briggs, et al. 2010. A draft sequence of the Neandertal genome. Science 328:710-722.

Guthrie, RD. 1970. Bison evolution and zoogeography in North America during the Pleistocene. The Quarterly review of biology 45:1-15.

Guthrie, RD. 1990. Frozen Fauna of the Mammoth Steppe: The story of Blue Babe. Chicago and London: The University of Chicago Press.

Halbert, ND. 2005. The utilization of genetic markers to resolve modern management issues in historic bison populations: implications for species conservation: Texas A&M University.

Ho, SYW, R Lanfear, L Bromham, MJ Phillips, J Soubrier, AG Rodrigo, A Cooper. 2011a. Time-dependent rates of molecular evolution. Molecular ecology 20:3087-3101.

Ho, SYW, R Lanfear, MJ Phillips, I Barnes, JA Thomas, SO Kolokotronis, B Shapiro. 2011b. Bayesian estimation of substitution rates from ancient DNA sequences with low information content. Systematic biology 60:366-375.

Ho, SYW, G Larson, CJ Edwards, TH Heupink, KE Lakin, PW Holland, B Shapiro. 2008a. Correlating Bayesian date estimates with climatic events and domestication using a bovine case study. Biology Letters 4:370-374.

Ho, SYW, MJ Phillips, A Cooper, AJ Drummond. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Molecular Biology and Evolution 22:1561-1568.

Ho, SYW, U Saarma, R Barnett, J Haile, B Shapiro. 2008b. The effect of inappropriate calibration: three case studies in molecular ecology. Plos One 3:e1615.

Janecek, LL, RL Honeycutt, RM Adkins, SK Davis. 1996. Mitochondrial Gene Sequences and the Molecular Systematics of the Artiodactyl Subfamily Bovinae. Molecular phylogenetics and evolution 6:107-119.

Keane, TM, CJ Creevey, MM Pentony, TJ Naughton, JO McLnerney. 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. BMC evolutionary biology 6:29.

Krause, J, Q Fu, JM Good, B Viola, MV Shunkov, AP Derevianko, S Paabo. 2010. The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. Nature 464:894-897.

Matukumalli, LK, CT Lawley, RD Schnabel, et al. 2009. Development and characterization of a high density SNP genotyping assay for cattle. Plos One 4:e5350.

McCue, ME, DL Bannasch, JL Petersen, et al. 2012. A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. PLoS genetics 8:e1002451.
McDonald, JN. 1980. North American bison. Their classification and evolution. Berkeley: University of California Press.
Miller, W, DI Drautz, A Ratan, et al. 2008. Sequencing the nuclear genome of the extinct woolly mammoth. Nature 456:387-390.
Minin, VN, EW Bloomquist, MA Suchard. 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Molecular Biology and Evolution 25:1459-1471.
Noonan, JP, M Hofreiter, D Smith, JR Priest, N Rohland, G Rabeder, J Krause, JC Detter, S Paabo, EM Rubin. 2005. Genomic sequencing of Pleistocene cave bears. Science 309:597-599.
Pertoldi, C, M Tokarska, JM Wójcik, et al. 2010. Phylogenetic relationships among the European and American bison and seven cattle breeds reconstructed using the BovineSNP50 Illumina Genotyping BeadChip. Acta Theriologica 55:97-108.
Prusak, B, G Grzybowski, G Zieba. 2004. Taxonomic position of Bison bison (Linnaeus 1758) and Bison bonasus (Linnaeus 1758) as determined by means of cytb gene sequence. Animal Science Papers and Reports 22(1):27-35.
Rambaut, A, AJ Drummond. 2007. Tracer v1.4, Available from beast.bio.ed.ac.uk/Tracer.
Rasmussen, M, Y Li, S Lindgreen, et al. 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 463:757-762.
Roe, EG. 1970. The North American buffalo: a critical study of the species in its wild state. Toronto,: University of Toronto Press,.
Shapiro, B, AJ Drummond, A Rambaut, et al. 2004. Rise and fall of the Beringian steppe bison. Science 306:1561-1565.
Shapiro, B, SYW Ho, AJ Drummond, MA Suchard, OG Pybus, A Rambaut. 2011. A Bayesian phylogenetic method to estimate unknown sequence ages. Molecular Biology and Evolution 28:879-887.
Skinner, M, O Kaisen. 1947. The fossil bison of Alaska and preliminary revision of the genus. Bulletin of the American Museum of Natural History 89.
Slatis, HM. 1960. An Analysis of Inbreeding in the European Bison. Genetics 45:275–287.
Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688-2690.
Stamatakis, A, P Hoover, J Rougemont. 2008. A rapid bootstrap algorithm for the RAxML Web servers. Systematic biology 57:758-771.
van Zyll de Jong, CG. 1986. A systematic study of recent bison, with particular consideration of the wood bison: National Museum of Natural Science Publication.
van Zyll de Jong, CG, C Gates, H Reynolds, W Olson. 1995. Phenotypic Variation in Remnant Populations of North American Bison. Journal of Mammalogy 76:391-405.
Vasil'ev, VA, EP Steklenev, EV Morozova, SK Semenova. 2002. DNA fingerprinting of individual species and intergeneric and interspecific hybrids of genera Bos and Bison, subfamily Bovinae. Genetika 38:515-520.
Verkaar, EL, IJ Nijman, M Beeke, E Hanekamp, JA Lenstra. 2004. Maternal and paternal lineages in cross-breeding bovine species. Has wisent a hybrid origin? Molecular Biology and Evolution 21:1165-1170.
Vonholdt, BM, JP Pollinger, KE Lohmueller, et al. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature 464:898-902.
Wang, Z, X Shen, B Liu, et al. 2010. Phylogeographical analyses of domestic and wild yaks based on mitochondrial DNA: new data and reappraisal. Journal of Biogeography 37:2332-2344.

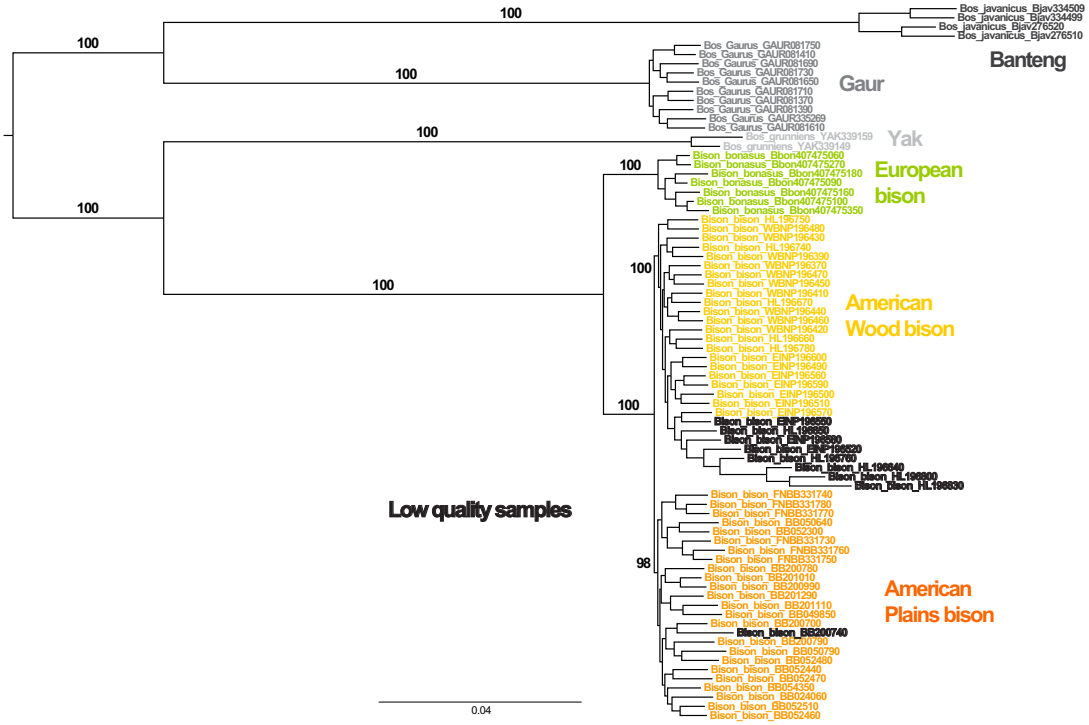
Supplementary Material:

Supplementary table 1. Details of the Steppe bison samples considered in the study and the resulting character composition from the selected 40,843 SNPs.

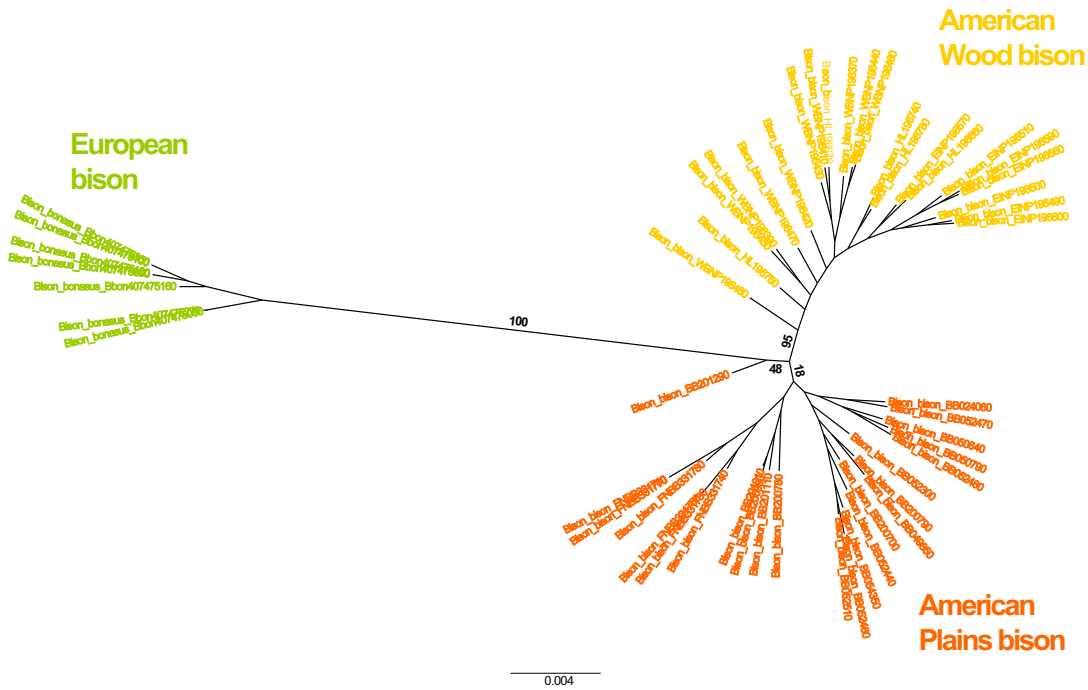
Sample ID	Library ID	Origin	AMS date	Amplification method	AA	AB	BB	Missing data	AB / (AA+AB+BB)
A002	**449476740**	Urals, Russia	51,800 ± 1,300[a]	sigma-WGA2	0.70	7.08	49.12	43.10	0.12
	449476741			sigma-WGA2	5.78	19.94	3.53	70.75	0.68
A3133	**449476720**	Yukon, Canada	26,360 ± 220[a]	sigma-WGA4	14.31	2.18	38.34	45.18	0.04
	449476721			sigma-WGA4	12.14	4.48	18.11	65.28	0.13
	449496520			Repli-g-FFPE	10.37	6.29	15.63	67.72	0.19
	449496560			sigma-WGA4	2.77	5.77	28.30	63.16	0.16
	449496580			sigma-WGA4	6.28	3.66	25.27	64.78	0.10
	449496570			sigma-WGA4	10.36	4.54	18.46	66.64	0.14
	449496800			sigma-WGA4	10.71	3.61	17.80	67.87	0.11
	449496780			sigma-WGA4	11.07	5.61	21.85	61.47	0.15
	449496790			sigma-WGA4	12.94	5.21	18.95	62.91	0.14
A5880	**449496840**	JiLin, China	~30,000[b]	sigma-WGA4	25.60	14.48	2.62	57.30	0.34
A5881	**449496850**	JiLin, China	~30,000[b]	sigma-WGA4	25.62	12.73	2.45	59.20	0.31
BS662	**449476730**	Alyoshkina Zaimka, Siberia, Russia	~20,000[b]	sigma-WGA2	9.90	5.97	37.07	47.07	0.11
	449476731			sigma-WGA2	3.44	12.26	22.06	62.23	0.32
	Decker, 2009			sigma-WGA2	3.26	18.11	24.39	54.23	0.40

[a] Uncalibrated radiocarbon dated years before present

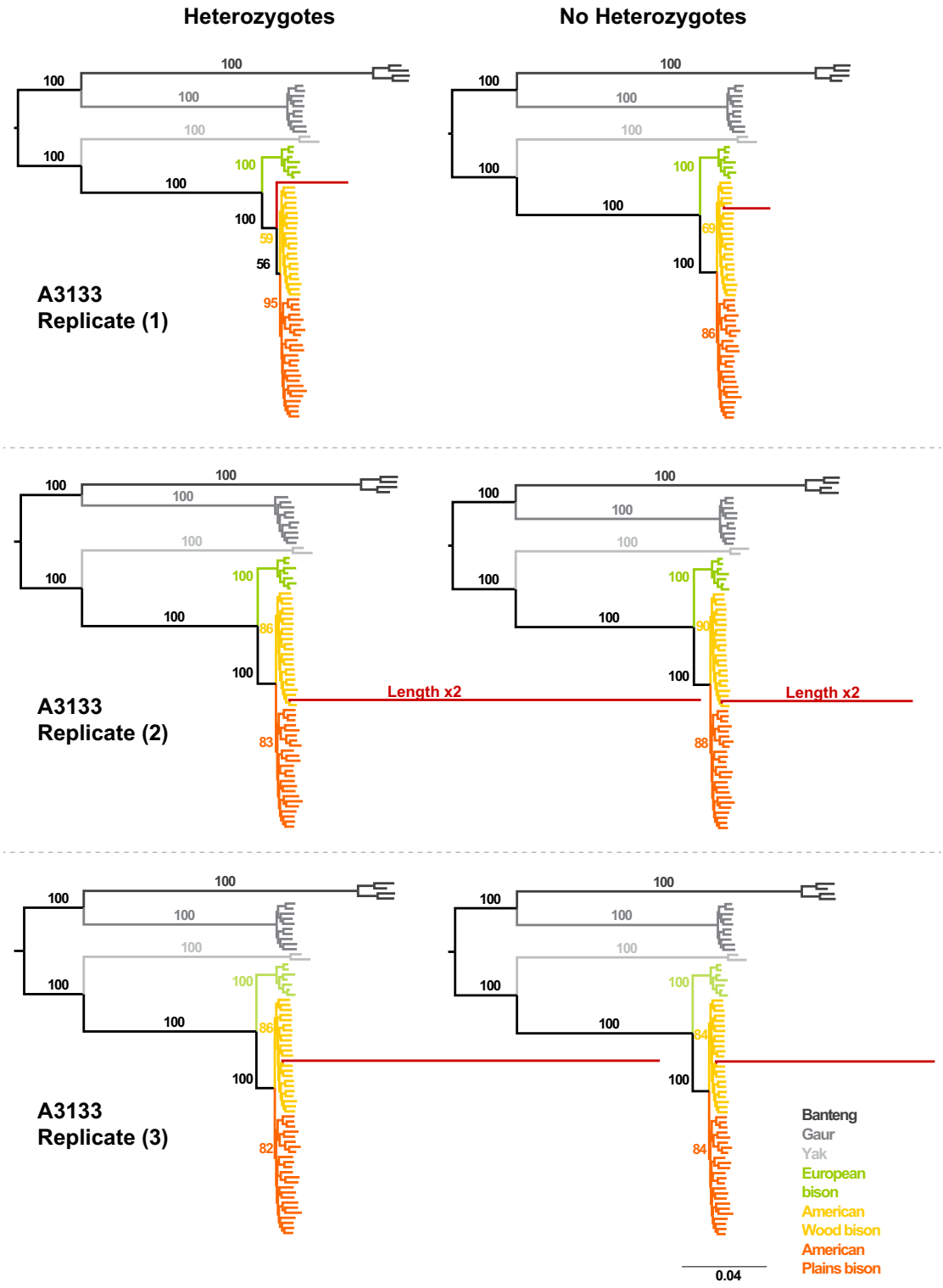
[b] Estimated age in years



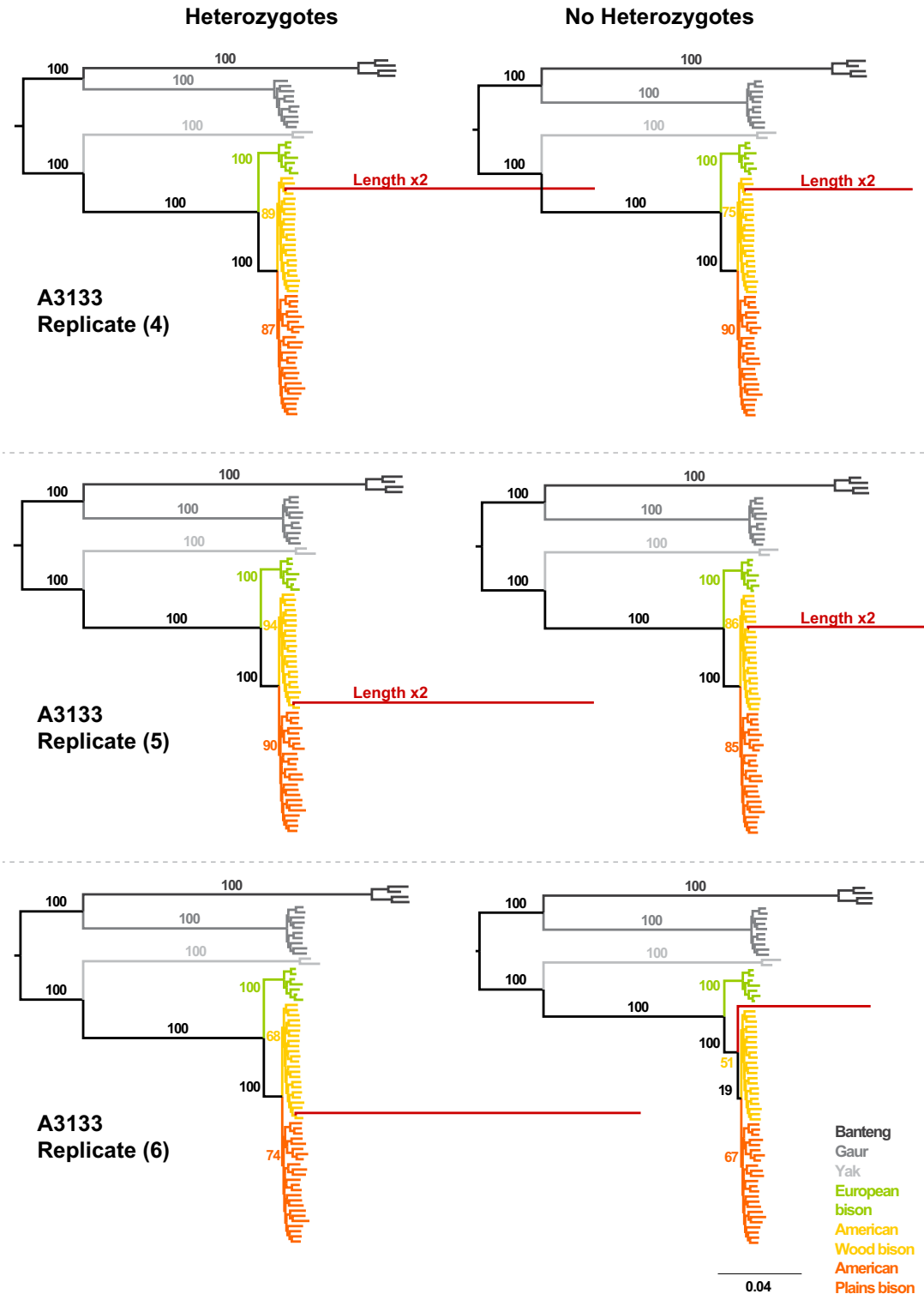
Supplementary Figure 1. ML phylogenetic tree showing the position of the low quality modern bison samples.



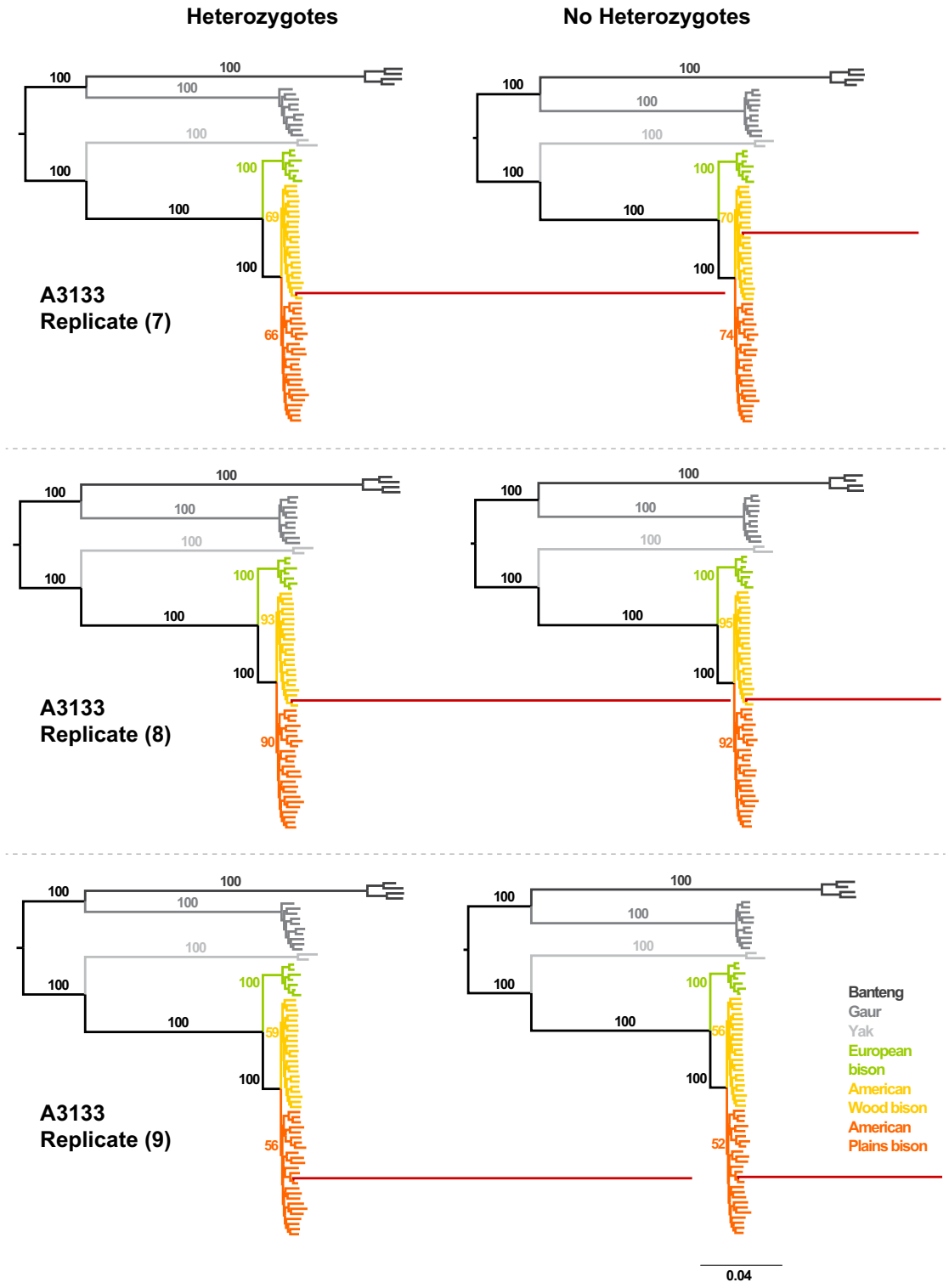
Supplementary Figure 2. ML phylogenetic tree of European and American bison calculated without heterozygote characters. Although the tree is very similar to the one calculated with heterozygotes (Figure 2b), reciprocal monophyly between Plains and Wood American bison is contradicated by the placement of a single individual BB201290.



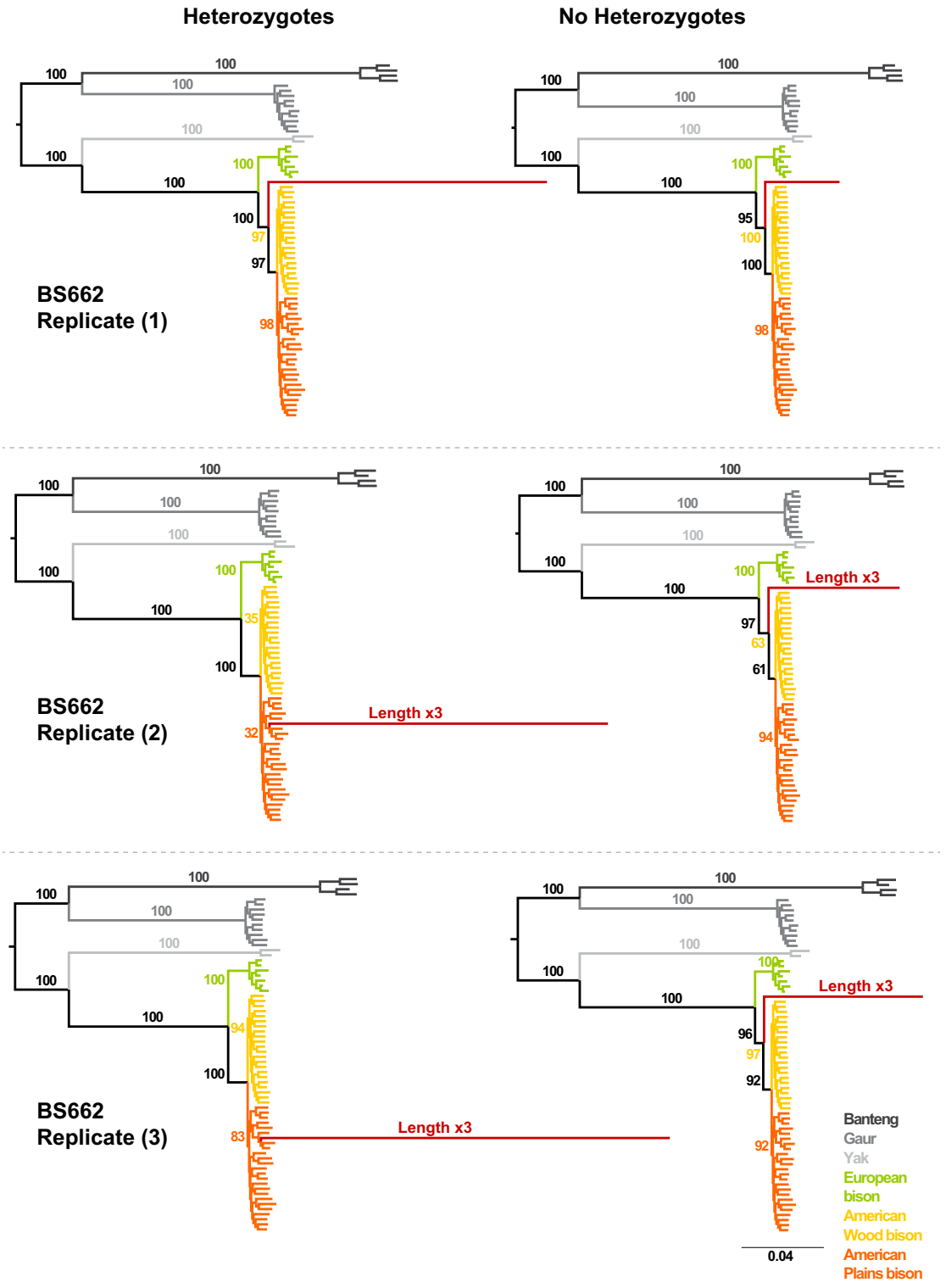
Supplementary Figure 3a. Individual phylogenetic placement of the library replicates 1 to 3 from the Yukon Steppe bison A3133 (red branch). All trees on the left side are calculated included heterozygote characters, while the heterozygotes were removed from the ancient libraries to infer the trees presented on the right side of the figure.



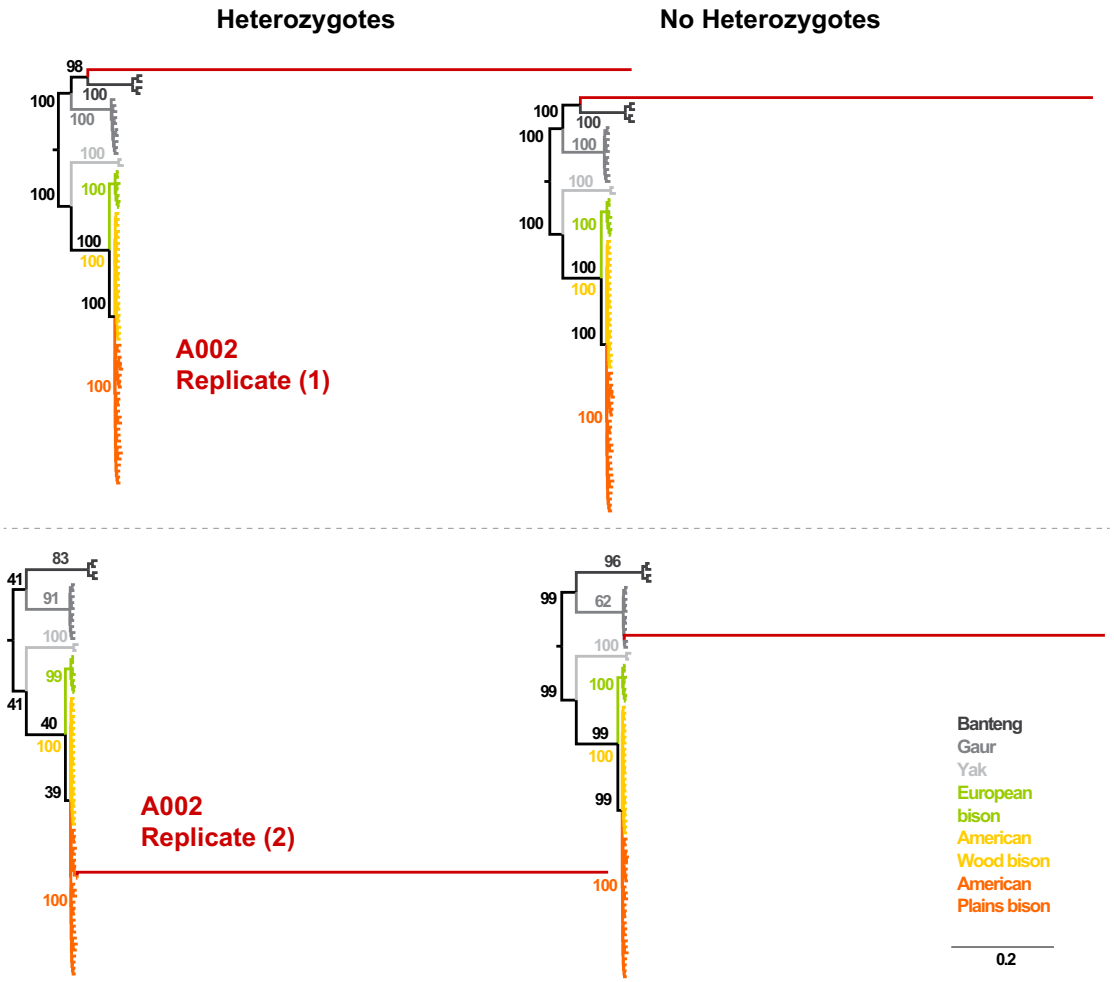
Supplementary Figure 3b. Individual phylogenetic placement of the library replicates 4 to 6 from the Yukon Steppe bison A3133.



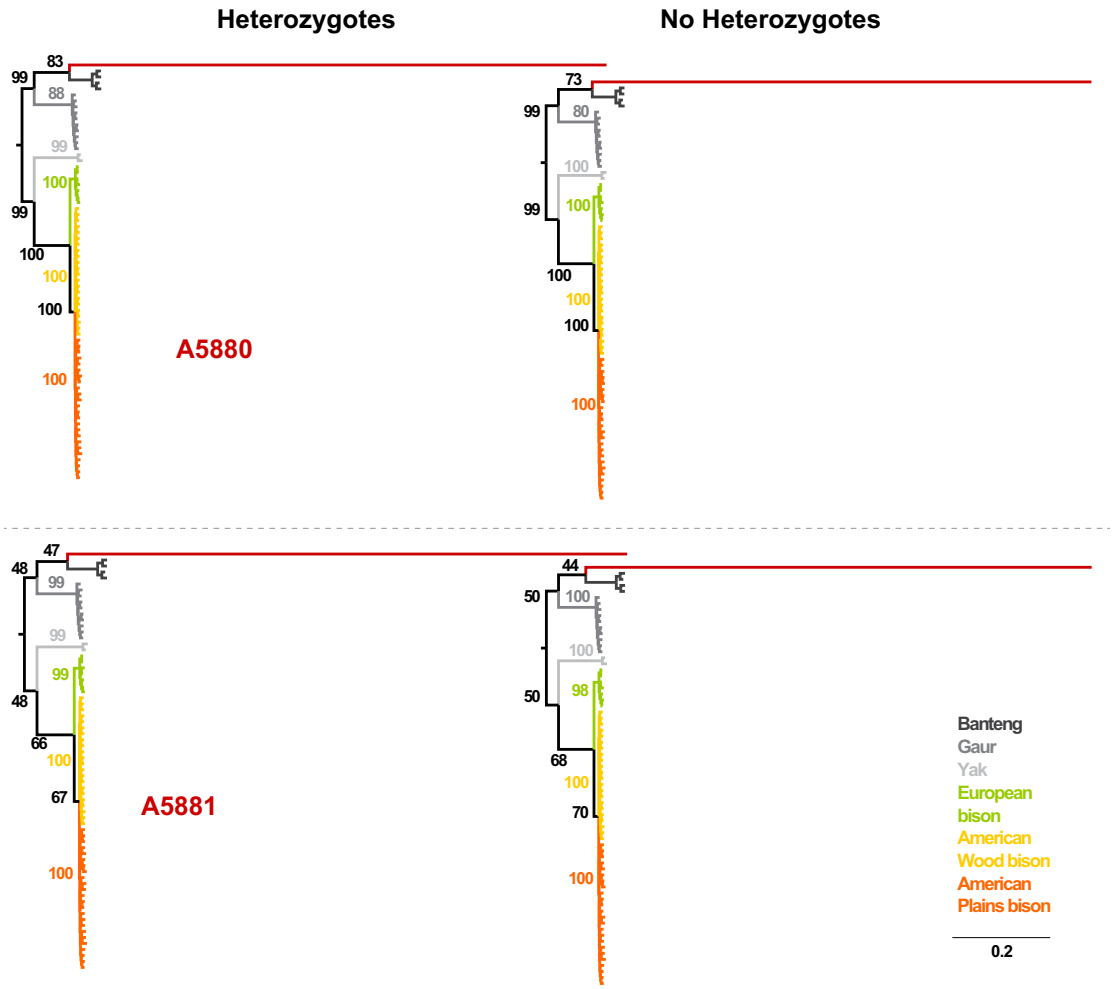
Supplementary Figure 3c. Individual phylogenetic placement of the library replicates 7 to 9 from the Yukon Steppe bison A3133.



Supplementary Figure 4. Individual phylogenetic placement of the library replicates 1 to 3 from the Siberian Steppe bison BS662.



Supplementary Figure 5. Individual phylogenetic placement of the two library replicates from the Urals Steppe bison A002.



Supplementary Figure 6. Individual phylogenetic placement of the library from the Chinese Steppe bison A5880 and A5881.

Chapter 4:

Paleoclimatic impacts on European bovid megafauna in the Late Pleistocene

Statement of authorship

Paleoclimatic impacts on European bovid megafauna in the Late Pleistocene

Julien Soubrier (Candidate)
Co-developed the research concept, performed the phylogenetic analyses, co-interpreted the results, developed the figures and wrote the manuscript.
I hereby certify that the statement of contribution is accurate.

Signed …… …… *Date* 02.04.12

Kefei Chen
Co-developed the research concept, processed the samples, performed ancient DNA research and library construction, co-produced the BovineSNP50 BeadChip genotypes, co-analysed the data, co-interpreted the results and co-wrote the manuscript.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… *Date* ……17/04/2012……

Beth Shapiro
Provided samples, performed initial experiments, co-interpreted the results and assisted with phylogenetic analyses.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… *Date* ……3 Apr 2012……

Simon Y.W. Ho
Co-interpreted the results, supervised phylogenetic analyses and manuscript writing.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… *Date* ……05.04.2012……

Wolfgang Haak
Collected samples and supervised manuscript writing.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… *Date* 02/04/2012

Michael S. Y. Lee
Co-developed the research concept, supervised the phylogenetic analyses and manuscript writing.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ……. …… *Date* ……17:4:12

Alan Cooper
Co-developed the research project, co-interpreted the results, co-wrote the manuscript and supervised the project.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ... ... *Date* ……05.04.2012……

Paleoclimatic impacts on European bovid megafauna in the Late Pleistocene

Julien Soubrier[1*], Kefei Chen[1*], Beth Shapiro[2], Simon Y.W. Ho[3], Wolfgang Haak[1], Michael S. Y. Lee [5] and Alan Cooper[1]

[1] Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, University of Adelaide, Adelaide, SA 5005, Australia;

[2] Department of Biology, The Pennsylvania State University, University Park, Pennsylvania, United States of America;

[3] School of Biological Sciences, University of Sydney, Sydney, NSW, Australia;

[4] Department of Anthropology, University of Arizona, Tucson, AZ 85721-0030; United States of America;

[5] South Australian Museum, North Terrace, Adelaide, South Australia 5000, Australia

[6] Division of Animal Sciences, University of Missouri, Columbia, MO 65211, USA.

[*] These authors contributed equally to this work

ABSTRACT

The fossil record contains a wide diversity of bison species and/or sub-species distributed across Eurasia and North-America during the late Pleistocene, although only two species remain alive today (the American bison, *Bison bison*, and the European bison, *Bison bonasus*). Recently, ancient DNA has been used to study the evolutionary history of North American and Beringian bison in the late Pleistocene, revealing dramatic population changes around the Last Glacial Maximum (LGM). In contrast, the genetic diversity of European bison through this time period is yet to be explored. Here, we extract and analyse ancient DNA from 48 European late Pleistocene bovid samples, collected from the North Sea, the Caucasus and the Urals. Phylogenetic analyses reveal a previously unknown species of bison, sister taxa to the European bison. In addition, the data record an unexpected series of rapid population replacements between the new species and Steppe bison across the three European sites dating back to beyond Oxygen Isotope Stage 3, ca. 55 kyr. The timing of the transitions between the species appear to closely correlate with known environmental changes, and are compared to a long-term paleoenvironmental record from southwestern Germany.

INTRODUCTION

The Late Pleistocene (126-11 kyr) climate record contains a series of large-scale oscillations with pronounced environmental effects (Martin 1984; Wolff et al. 2010). The role of these environmental changes as a driver of the widespread extinction of mammalian megafauna (>42kg) around the world during this time is widely debated (Martin 1984; Graham et al. 1996; Stuart, Lister 2007). A key issue is how to separate these impacts from those caused by the arrival of human populations, which were undergoing a contemporaneous expansion during the latter part of the Pleistocene (50-11 kyr), notably during warming phases. In islands and isolated continents such as Australia, the correlation between human colonisation and megafaunal extinctions is pronounced but in Eurasia, North and South America the nature and extent of these impacts, and the degree to which they were synergistic remains unresolved (Graham et al. 1996; Barnosky et al. 2004; Barnosky, Lindsey 2010; Barnosky et al. 2011; Lorenzen et al. 2011; Rule et al. 2012; Stewart, Stringer 2012).

Europe is an ideal situation to examine the timing and nature of megafaunal extinctions. Large radiocarbon dating surveys of fossil mammal bones are available (Stuart, Lister 2007), and multiple cave environments exist which are suitable for DNA preservation. Detailed climate records are available, ranging from Greenland ice cores (GISP2) to lake and pollen records. Recently, a detailed paleovegetation reconstruction spanning Marine Isotope Stages (MIS) 4-1 (MIS 4, 71-60ky; MIS 3, 60-24 ky; MIS 2, 24-11 ky; MIS 1, 11-0 ky) for western Europe has became available through a combined analysis of sediments from Maar lakes in the Eifel region of SW Germany (Sirocko et al. submitted). Ancient mitochondrial DNA (mtDNA) studies have identified large-scale changes in the genetic diversity of ancient European megafaunal populations over time, ranging from cave lions (Barnett et al. 2009), cave bears (Orlando et al. 2002; Stiller et al. 2010), cave hyenas (Rohland et al. 2005), horses (Lorenzen et al. 2011), mammoths (Barnes et al. 2007; Nystrom et al. 2010), saiga antelope (Campos et al. 2010), bison (Shapiro et al. 2004) and Neandertals (Dalén et al. 2012). However, to date there has been no detailed long-term, multi-species analysis of the association between the timing of extinction events and paleoenvironmental changes.

Late Pleistocene bovid fossils are one of the most complete European megafaunal records, with several species currently recognised. These include the extinct aurochs (*Bos primigenius*) and three species of bison: the extinct Steppe bison (*Bison priscus*) which ranged from Alaska to western Europe; the extinct European Pleistocene short-horned forest bison (*Bison schoetensacki*), and the living but endangered wisent, or European bison (*Bison bonasus*), whose fossils appear around the Pleistocene/Holocene boundary (ca. 11ky) or later. The fossil record is somewhat complicated by the similarity between the bison species, and the difficulty in taxonomically identifying bovid post-cranial elements.

These taxa have widely varying histories. The aurochs is generally accepted as the ancestor of modern cattle, which were independently domesticated from ancient populations of aurochs in different parts of the world. MtDNA evidence suggests that Late Pleistocene aurochs populations in northern Europe did not contribute to European cattle (*Bos taurus*). In contrast, populations in southern Europe did share mtDNA lineages with modern *Bos taurus*, although it is not clear whether this was due to a direct genetic contribution to the domestication process, or common shared ancestry (Lari et al. 2011). The shared mtDNA lineages appear in southern European aurochs specimens from the Late Pleistocene, before the appearance of domestic cows and therefore cannot result from interbreeding between wild and domestic populations. The history of the aurochs has been well studied genetically (Beja-Pereira et al. 2006; Edwards et al. 2007; Edwards et al. 2010; Lari et al. 2011; Bollongino et al. 2012), and instead this study focuses on the bison, which have not been examined in detail.

Ancient DNA studies revealed that Late Pleistocene populations of Steppe bison in the Old World constitute a small subset of the total mitochondrial DNA (mtDNA) diversity observed in eastern Beringia, *i.e.* Alaska and Yukon (Shapiro et al. 2004) at this time. This is surprising, given that the fossil record indicates Steppe bison originated in Asia, and only recently colonised the New World, ca. 200-400 ky (McDonald 1980). To explain this, it has been suggested that former populations of the Old World underwent a major extinction phase(s), or were replaced by the re-invading Alaskan population by some other means, prior to the earliest mtDNA sequences (ca. 80 kyr) (Shapiro et al. 2004). The last Steppe bison in the Old World are thought to have become extinct in the late Holocene, although the exact date is unknown (Lazarev, Boeskorov, Tomskaya 1998).

All living European bison, *Bison bonasus*, are descended from 17 animals originating from two small late-19th/early-20th century populations (Slatis 1960). Extant populations now exceed more than 2000 individuals, of which all pure-bred European bison represent the recombination of only 12 diploid sets of genes (Slatis 1960). Separate Caucasus and Carpathian subspecies of the European bison (*B. b. caucasicus and B. b. hungarorum*) are recognised from the fossil record, but became extinct in the wild in 1927 and 1790, respectively. Surviving populations continue to be threatened by habitat destruction, inbreeding, disease and human hunting. The recent population bottleneck in European bison has made it difficult to reconstruct their genetic history. While morphological similarities, autosomal phylogenies (Chapter 3), and the ability to interbreed to produce fertile female offspring, each suggest a close evolutionary relationship with American bison (*Bison bison*), mitochondrial phylogenies depict a very different evolutionary history, in which European bison are actually more closely related to cattle than to American bison (Janecek et al. 1996; Verkaar et al. 2004). Two hypotheses have been proposed to explain the difference between

the phylogenies (Verkaar et al. 2004): lineage sorting, in which two distinct mitochondrial lineages survived in the bison/yak lineage until the recent species-level split with European bison, or sex-biased genetic introgression, in which steppe bison males repeatedly mated with an ancestral bovid in the ox/zebu lineage, resulting in the sudden appearance of a new species with bison-like morphology and autosomal DNA, but ox/zebu-like mitochondrial DNA.

The paucity of early, clearly identifiable *B. bonasus* remains further complicates this issue. The oldest *B. bonasus* are from the Late Pleistocene/early Holocene (Flerov 1979; Kahlke 1999), or even the late Holocene (Pucek 1986), and little agreement about ancestral forms has been reached (see Stuart 1991; Kahlke 1999; Bauer 2001). It is thought that the direct ancestor of the European bison is the Pleistocene forest bison *B. schoetensacki* (Kurtén 1968; Geist, Karsten 1977), although some have suggested *B. priscus* (Flerov 1979; Bauer 2001).

To clarify the evolutionary history of bovids in Europe, and investigate the potential impacts of environmental change in megafaunal extinctions, fossil remains of bison were collected from four separate locations across Europe (the Caucasus, the North Sea, the Urals and Austria), covering a period of more than 60 ky.

METHOD

Data

New samples

A total of 68 late Pleistocene samples were collected from four regions across the full geographic range of Europe (Table 1). The main sample was from NE Europe, and represents isolated bones excavated from a wide variety of cave deposits throughout the Urals Mountains and surrounding areas, including river deposits, held in collections at the Zoological Museum of the Institute of Plant and Animal Ecology (ZMIPAE) in Ekaterinburg, Russia. In Western Europe, a second collection of bones was analysed from late Pleistocene deposits on the North Sea bed. These specimens were recovered by trawling operations and have little stratigraphic information, and are curated by the North Sea Network (NSN) in the Netherlands. In SE Europe, bovid bone fragments from excavations at Mezmaiskaya Cave in the Caucasus Mountains were analysed. This high altitude site contains both Neandertal and early *H. sapiens* remains (Skinner et al. 2005), and samples were obtained from the Laboratory of Prehistory, in St Petersburg. The last site consisted of a small number of bones from central European Holocene sites, held in the collections of the Natural History Museum, Vienna (VNHM). In all cases, samples were from bones identified as bovid post-cranial samples as cranial material is rare for this time period.

Table 1. List of all samples from Urals, North Sea, Caucasus and Austria used in the study.

	Sample ID	Sequence > 400bp	Species	AMS date		Calibrated dates			Origin	Type
				Oxdate	Oxerr	Mean	Low	High		
	A001	✓	BisonX	12565	55	14773	14240	15159	Rasik 1 (ZMIPAE)	Pelvis fragment
	A002	✓	*B. priscus*	51800	1300	52145	49495	55047	Sur'ya 5 (ZMIPAE)	Metacarpal
	A003	✓	BisonX	12505	55	14655	14206	15072	Voronovka (ZMIPAE)	Humerus
	A004	✓	BisonX	19010	80	22698	22320	23269	Rasik 1 (ZMIPAE)	Metacarpal
	A005	✓	BisonX	15310	70	18563	18161	18764	Ladeinyi Kamen (ZMIPAE)	Femur
	A006	✓	BisonX	18880	90	22536	22210	23235	Sur'ya 5 (ZMIPAE)	Metatarsal
	A007	✓	BisonX	58300	2900	60660	53641	70691	Sur'ya 3 (ZMIPAE)	Metatarsal
	A008	✓	*B. priscus*	31560	210	35960	35285	36582	Dinamitnaya (ZMIPAE)	Metacarpal
	A011	✓	BisonX	60900	**INF**	79077	50088	111451	Sur'ya 5 (ZMIPAE)	Metatarsal
	A012	✗	Contamination						Sur'ya 5 (ZMIPAE)	Metacarpal
	A013	✓	*B. priscus*	48400	900	48558	46760	50464	Rasik 1 (ZMIPAE)	Tibia
	A014	✓	*B. priscus*	33820	260	38673	37675	39451	Bobylek (ZMIPAE)	Tibia
Urals	A015	✗	Contamination						Yurovsk (ZMIPAE)	Femur
	A016	✓	BisonX	49600	1200	49890	47459	52529	Gofmana (ZMIPAE)	Humerus
	A017	✓	BisonX	18850	90	22495	22184	23229	Sur'ya 5 (ZMIPAE)	Upper mandible
	A018	✓	BisonX	13120	60	15907	15241	16482	Sur'ya 5 (ZMIPAE)	Radius
	BS588	✓	*B. bonasus*	16810	65	19964	19584	20271	Sur'ya 5 (ZMIPAE)	Metapodial
	BS592	AY748756	*B. priscus*	42500	450	45725	44975	46500	Chernye Kosti (ZMIPAE)	Femur
	BS599	✓	BisonX	26330	120	31003	30759	31215	Kholodnyi (ZMIPAE)	Tibia
	BS604	✓	BisonX	55400	1800	56123	52281	60568	Sur'ya 5 (ZMIPAE)	Astralagus
	BS606	✓	BisonX	25000	100	29881	29522	30221	Kholodnyi (ZMIPAE)	Bone fragment
	BS660	AY748766	*B. priscus*	29500	140	34195	33610	34655	Sur'ya 5 (ZMIPAE)	Metapodial
	BS674	AY748775	*B. priscus*	29060	140	33765	33220	34469	Kholodnyi (ZMIPAE)	Phalanx
	BS708	AY748793	*B. priscus*	47050	750	47158	45665	48725	Rasik 1 (ZMIPAE)	Femur
	BS713	AY748795	*B. priscus*	30970	180	35595	34957	36283	Irtysh River (ZMIPAE)	Metatarsal
	A2791	✓	BisonX	53800	**INF**	74612	50092	107567	North Sea bed deposit (NSN)	-
	A2792	✓	*B. priscus*	29100	150	33811	33251	34485	North Sea bed deposit (NSN)	-
	A2793	✓	*B. priscus*	28340	130	32629	32003	33185	North Sea bed deposit (NSN)	-
	A2794	✗							North Sea bed deposit (NSN)	-
	A2795	✓	BisonX	29010	160	33708	33142	34469	North Sea bed deposit (NSN)	-
	A2796	✓	*B. priscus*	43850	650	47125	45672	48890	North Sea bed deposit (NSN)	-
	A2797	✓	*B. taurus*						North Sea bed deposit (NSN)	-
	A2798	✓	BisonX	29230	150	33936	33365	34530	North Sea bed deposit (NSN)	-
	A2799	✓	*B. taurus*						North Sea bed deposit (NSN)	-
	A2800	✓	Elk						North Sea bed deposit (NSN)	-
North Sea	A2801	✗	Contamination						North Sea bed deposit (NSN)	-
	A2802	✗							North Sea bed deposit (NSN)	-
	A2803	✗							North Sea bed deposit (NSN)	-
	A2804	✗							North Sea bed deposit (NSN)	-
	A2805	✗							North Sea bed deposit (NSN)	-
	A2806	✗							North Sea bed deposit (NSN)	-
	A2807	✗							North Sea bed deposit (NSN)	-
	A2808	✓	BisonX	61500	**INF**	55979	50038	67952	North Sea bed deposit (NSN)	-
	A2809	✓	BisonX	61300	**INF**	58581	50016	75258	North Sea bed deposit (NSN)	-
	A2810	✓	*B. taurus*						North Sea bed deposit (NSN)	-
	A2811	✓	BisonX	62000	**INF**	57640	50035	71248	North Sea bed deposit (NSN)	-
	A4081	✓	BisonX	Not dated					Mezmaiskaya, level 3	Long Bone
	A4082	✓	BisonX	Not dated					Mezmaiskaya, level 3	Long Bone
	A4083	✓	BisonX	Not dated					Mezmaiskaya, level 3	Long Bone
	A4084	✓	BisonX	Not dated					Mezmaiskaya, level 3	Long Bone
	A4085	✓	BisonX	Not dated					Mezmaiskaya, level 3	Long Bone
	A4086	✗							Mezmaiskaya, level 3	Long Bone
	A4087	✓	BisonX	Not dated					Mezmaiskaya, level 3	Long Bone
	A4088	✓	BisonX	Not dated					Mezmaiskaya, level 3	Long Bone
	A4089	✓	BisonX	59400	**INF**	96533	50080	128118	Mezmaiskaya, level 2B4	Long Bone
	A4090	✓	*B. priscus*	59400	**INF**				Mezmaiskaya, level 2B4	Long Bone
	A4091	✓	BisonX	59700	**INF**	77885	50037	119808	Mezmaiskaya, level 2B4	Long Bone
Caucasus	A4092	✓	BisonX	56600	**INF**	54627	50020	65094	Mezmaiskaya, level 2B4	Long Bone
	A4093	✓	*B. bonasus*	56300	**INF**				Mezmaiskaya, level 2B3	Long Bone
	A4094	✓	BisonX	56500	**INF**	53962	50010	62790	Mezmaiskaya, level 2B3	Long Bone
	A4095	✗							Mezmaiskaya, level 2B2	Long Bone
	A4096	✗							Mezmaiskaya, level 2B2	Long Bone
	A4097	✗							Mezmaiskaya, level 2A	Long Bone
	A4098	✓	Brown bear						Mezmaiskaya, level 2A	Long Bone
	A4099	✗							Mezmaiskaya, level 2A	Long Bone
	A4100	✗							Mezmaiskaya, level 2A	Long Bone
	A4101	✗							Mezmaiskaya, level 1C	Long Bone
	A4102	✗							Mezmaiskaya, level 1C	Long Bone
	A4103	✓	*B. taurus*						Mezmaiskaya, level 1-2	Long Bone
	A4104	✓	BisonX	12160	40	14008	13852	14163	Mezmaiskaya, level 1-2	Long Bone
	BS593	✓	*B. bonasus*	5090	60	5824	5662	5982	Steiermark (VNHM)	Femur
Austria	BS600	✓	*B. bonasus*	3430	50	3696	3571	3833	Steiermark (VNHM)	Femur
	BS607	✓	*B. bonasus*	1370	50	1287	1179	1371	Oberosterreich (VNHM)	Femur

Published control region sequences

To provide comparative data, 302 published mitochondrial control region sequences of 628 bp were assembled (Supplementary Table 1), representing the following bovid mitochondrial lineages: European bison (*Bison bonasus*), American bison (*Bison bison*), Steppe bison (*Bison priscus*), zebu (*Bos indicus*), and cattle (*Bos taurus*). Among these published sequences, 5 sequences were from Steppe bison collected in the Urals that had been analysed in an earlier study (Table 1).

Published whole mitochondrial genomes

Whole mitochondrial genome sequences were also available for 28 bovid specimens: 2 European bison (*Bison bonasus*), 8 American bison (*Bison bison*), 4 yak (*Bos grunniens*), 3 zebu (*Bos indicus*), 8 cattle (*Bos taurus*) and 3 buffalo (*Bubalus bubalis*) (Supplementary Table 2). These were assembled and analysed to generate a phylogenetic topology that could be compared to the shorter control region dataset.

Ancient DNA Extraction and Amplification

Ancient DNA was extracted from powdered bone using standard phenol/chloroform/centrifugal filtration methods (Shapiro et al. 2004). DNA extractions and PCR reactions were performed and set-up accordingly in a geographically isolated, authentic ancient DNA laboratory at the University of Adelaide. A ~600 base-pair (bp) fragment of the mitochondrial control region was amplified in one to four (overlapping) fragments, depending on the quality of the specimen. Two-step multiplex PCR amplifications were performed using primers designed for the bovid mitochondrial control region. Multiplex primer sets A and B were set up separately (supplementary Table 3). Multiplex PCR was performed in a final volume of 25 µl containing 2 µl of aDNA extract, 1 mg/ml rabbit serum albumin (RSA; Sigma, fraction V), 6 mM MgSO4, 0.2 µM of each primer, 500 µM of each dNTP, 2 U Platinum *Taq* Hi-Fidelity and 1 × PCR buffer (Invitrogen Ltd., UK). Multiplex PCR conditions were initial denaturation at 95 °C for 2 min, followed by 35 cycles of 94 °C for 15 sec, 55 °C for 20 sec and 68 °C for 30 sec, and a final extension at 68 °C for 10 min at the end of the 35 cycles. Multiplex PCR products were then diluted to 1:10 as template for the second step of simplex PCR. The second step simplex PCR using Amplitaq Gold (Applied Biosystems) or Hotmaster™ *Taq* DNA polymerase (5Prime, Milton, Qld) was conducted in a final volume of 25 µl containing 1 µl of diluted multiplex PCR product, 2.5 mM MgCl2, 0.4 µM of each primer, 200 µM of each dNTP, 1 U Amplitaq Gold/ Hotmaster *Taq* polymerase and 1 × PCR buffer. The second step simplex PCR conditions were initial denaturation at 95 °C for 2 min, followed by 35 cycles of 94 °C for 20 sec, 55 °C for 15 sec and 72 °C for 30

sec, and a final extension at 72 °C for 10 min at the end of the 35 cycles. Multiple PCR fragments were cloned to evaluate the extent of DNA damage and to detect the presence of nuclear mitochondrial inserts (numts). Multiple samples were also independently replicated, having been originally extracted and sequenced at the Henry Wellcome Ancient Biomolecules Centre at Oxford University.

One-step simplex PCR amplifications using Platinum *Taq* Hi-Fidelity polymerase were performed on a DNA Engine Tetrad2 Peltier Thermal Cycler (Bio-Rad) in a final volume of 25 µl containing 1 µl of aDNA extract, 1mg/ml rabbit serum albumin (RSA; Sigma, fraction V), 2 mM MgSO4, 0.6 µM of each primer, 250 µM of each dNTP, 1.25 U Platinum *Taq* Hi-Fidelity and 1 × PCR buffer (Invitrogen Ltd., UK). The conditions of PCR amplification were initial denaturation at 95 °C for 2 min, followed by 50 cycles of 94 °C for 20 sec, 55 °C for 20 sec and 68 °C for 30 sec, and a final extension at 68 °C for 10 min at the end of the 50 cycles. Negative extraction controls as well as non-template PCR controls were used throughout all experiments. PCR products were then checked by electrophoresis on 3.5-4.0% agarose TBE gels, and visualized after ethidium bromide staining on an UV transilluminator. PCR amplicons were purified using AMPure magnetic beads (Agencourt®, Beckman Coulter) according to manufacturer's instruction.

Sequencing

All purified PCR products were bi-directionally sequenced with the ABI Prism® BigDye™ Terminator Cycle Sequencing Kit version 3.1 (Applied Biosystems). The sequencing reactions were performed in a final volume of 10 µl containing 3.2 pmol of primer, 0.25 µl Bigdye terminator premixture, 1.875 µl of 5 × sequencing buffer. The reaction conditions contained initial denaturation at 95 °C for 2 min, 25 cycles with 95 °C for 10 sec, 55 °C for 15 sec, 60 °C for 2 min 30 sec. Sequencing products were purified using Cleanseq magnetic beads (Agencourt®, Beckman Coulter) according to the manufacturer's protocol. All sequencing reactions were analyzed on an ABI 3130 DNA capillary sequencer (Applied Biosystems, Foster, CA).

Radiocarbon dating

All samples from which bison mitochondrial control region sequences were successfully amplified were sent for accelerator mass spectrometry (AMS) radiocarbon dating (except for seven samples from level 3 of the Mezmaiskaya cave, for which dates were expected to be older than AMS dating capabilities). The dating was performed by AMS facility at the Oxford Radiocarbon Accelerator Unit in the University of Oxford, the INSTAAR Laboratory for AMS Radiocarbon Preparation and Research (NSRL) in the

University of Colorado at Boulder, and Keck-Carbon Cycle AMS facility (KCCAMS) in the University of California, Irvine. The calibration of radiocarbon dates was performed using OxCal v4.1 and the IntCal09 curve (Reimer et al. 2009). The dates mentioned throughout the paper are kcal yr BP unless otherwise stated.

Phylogenetic analysis

Genetic identification of the new specimens

Sequences from two individuals did not match bovid haplotypes, and were identified as Brown bear and Elk by BLAST (see Table 1 and Discussion). This is presumably due to the elements being morphologically non-diagnostic.

To taxonomically identify the other bovid samples, a phylogenetic tree including the remaining new sequences and 302 published bovid sequences was calculated using both maximum likelihood and Bayesian methods: the substitution model HKY + G6 was selected through comparison of Bayesian Information Criterion scores in ModelGenerator v0.85 (Keane et al. 2006). A maximum likelihood analysis was performed with the program PhyML v3 (Guindon et al. 2010), using the best of NNI and SPR moves for tree topology searching and aLRT to establish the statistical support of internal branches. Two independent Bayesian analyses were performed using the program MrBayes v3.2.1 (Ronquist et al. 2012), with four chains of 10 million generations, sampled every 1000 generations. The first 50 % of each chain was discarded as burn-in before the majority-rule consensus tree was calculated.

The same method was applied to re-calculate the corresponding bovid phylogeny form 28 whole mitochondrial genomes, using HKY+I as model of evolution.

Dated phylogeny

Using the program BEAST v1.6.2 (Drummond, Rambaut 2007), a calibrated phylogenetic analysis based on the coalescent model was performed using all carbon dated samples identified as BisonX. The GMRF Skyride model (Minin, Bloomquist, Suchard 2008) was used to account for variations in the population demographic history of BisonX, and a strict clock was assumed. Replicated analyses, using a relaxed uncorrelated lognormal clock to account for potential rate variations, could not reject the hypothesis of a strict molecular clock for that data set. Calibrated radiocarbon dates associated with all sequences were used as calibration points. Some samples appear to be older than 55 ky, reaching the limit of AMS dating methods (one from the Urals, four from the North Sea and four from the Caucasus, see Table 1). As these dates have infinite error margins, the Tip Date Sampling option (Shapiro et al. 2011) implemented in BEAST was used to account for their age uncertainty. The dated samples from Mezmaiskaya Cave are from stratigraphic layer 2B, which lies on top of layer

3-4, and the latter has been dated at a maximum age of 78 ky BP (including error margins). Consequently, a lognormal distribution prior was set for each Caucasian sample date with a minimum bound of 40 ky and a 95 % of the prior probability less than 80 ky. A similar distribution was used for the five remaining samples with infinite date margins, using a 95 % prior probability lower than 150 ky. All parameters showed sufficient sampling (indicated by effective sample sizes above 200) after 50,000,000 iterations and 10% burn-in. In addition, a 'date randomization test' was conducted to check whether the temporal signal from the radiocarbon dates associated with the ancient sequences was sufficient to calibrate the tree (Ho et al. 2011). This test randomizes all dates associated with the sequences and repeats the phylogenetic analysis to ascertain whether the inferred mean rate of the randomized analysis is significantly different from the rate calculated using the correctly associated radiocarbon dates and sequences. This means that the 95% HPD of the randomized analysis should not overlap with the mean rate estimated without randomization.

Survey of temperature and paleo-vegetation record contemporaneous to BisonX samples

The calibrated AMS radiocarbon dates and posterior distribution of tip dates from the Tip Date Sampling analysis are compared with temperature variations indicated by $\delta^{18}O$ values from the North Greenland Ice Core Project (NGRIP, Wolff et al. 2010). The paleovegetation reconstructions for the Eifel region of SW Germany are also reported on the same graph (Sirocko et al. 2012, submitted).

RESULTS AND DISCUSSION

Ancient DNA mitochondrial typing

Mitochondrial control region sequences (>400bp) were successfully amplified from 50 out of 68 samples analysed. Three samples produced a mixture of cattle and bison amplification products, and were identified as contaminated and removed from all analyses. In contrast, four samples produced consistent *Bos taurus* sequences: one Caucasus sample and three North Sea samples. Repetition and cloning experiments suggested that these are genuine, and the haplotypes were rare or previously unknown (Figure 1).

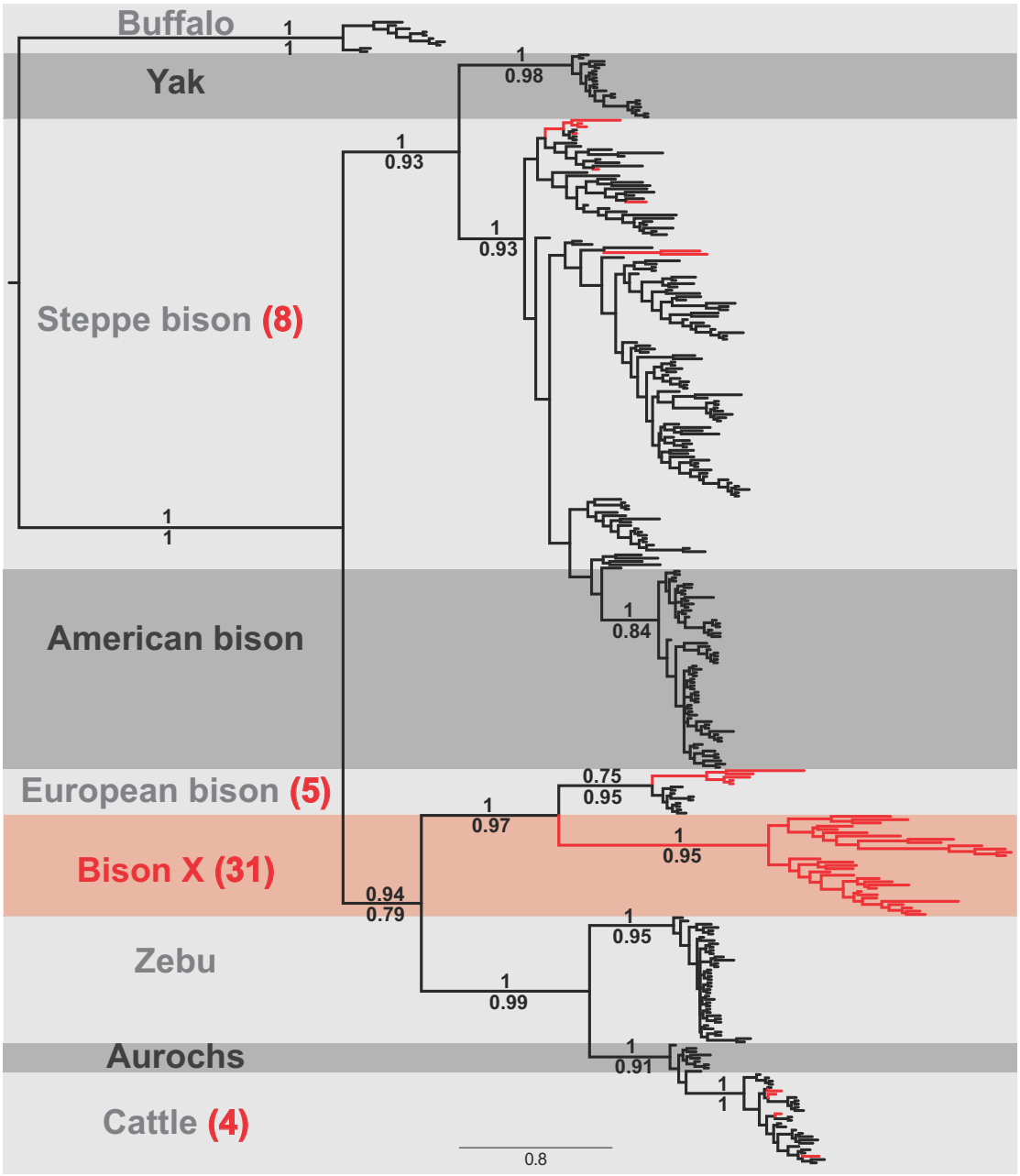


Figure 1. Phylogenetic tree of 350 bovid D-loop sequences. The position of the 48 newly sequenced individuals are marked in red. Numbers above branches represent posterior probabilities from MrBayes, and while those underneath represent aLRT support values from PhyML. The deep split between European bison and BisonX is obvious, and is of comparable scale to yak and Steppe/American bison.

Position of new samples in the bovid mitochondrial phylogeny

The 48 new bovid control region sequences were included in a larger alignment of 350 bovid sequences and phylogenetically analysed (Figure 1). The resulting tree allowed the genetic characterization of five European bison (*Bison bonasus*) mitochondrial haplotypes, four cows (*Bos taurus*), eight Steppe bison (*Bison priscus*) and 31 individuals belonging to an unknown clade. The unidentified bison (we will refer to them as BisonX) form a monophyletic clade sister to European bison. This new clade has high statistical support from both ML and Bayesian calculations, and a deep genetic split is observed between European bison and BisonX, indeed deeper than the one observed between yak and Steppe/American bison. Such a deep genetic split is likely to reflect a species level separation, although this is difficult to test in an extinct species.

The current study confirms current concepts that the European bison (*Bison bonasus*) was rare in Europe prior to the Holocene, raising the possibility that BisonX might represent a diverse but ancestral *B. bonasus* lineage, or perhaps the short-horned Pleistocene European forest wisent *Bison schoetensacki*, a putative ancestor of *B. bonasus*. However, ancient *B. bonasus* specimens were identified from the Urals at 20 ky BP, and in Mezmaiskaya cave in the Caucasus with an infinite date >56.3 ky BP. The latter sample is the oldest European bison specimen described, and demonstrates that BisonX and European bison were contemporaneous in the Caucasus around 50-60Kyrs BP. Furthermore, while the ancient sequences fall outside of the modern mtDNA variation in European bison, the clade remains well supported, and clearly distinct from BisonX.

A key issue is that the group formed by European bison and BisonX appears to be closely related to the zebu, aurochs and cattle rather than the other Bison (Figure 1). This topology differs from the close morphological and nuclear genomic relationship between European and American bison (see Chapter 3). The relationship is not due to the 31 new BisonX sequences as previous analyses of short mitochondrial sequences have also found a grouping of European bison with cattle/zebu (Janecek et al. 1996; Verkaar et al. 2004). It is difficult to assess whether the short control region fragment has enough signal to correctly reconstruct a phylogeny with all *Bos* and *Bison* species on the same tree (although ML and Bayesian statistics strongly support the robustness of the inferred topology). To confirm the contrasting mitochondrial and nuclear topologies, we analysed 28 whole mitochondrial genome sequences from American bison, European bison, cattle, zebu, yak and buffalo (Figure 2). The resulting phylogeny shows strong support for the placement of European bison as sister taxa to Zebu/Cattle, confirming the validity of the control region phylogeny.

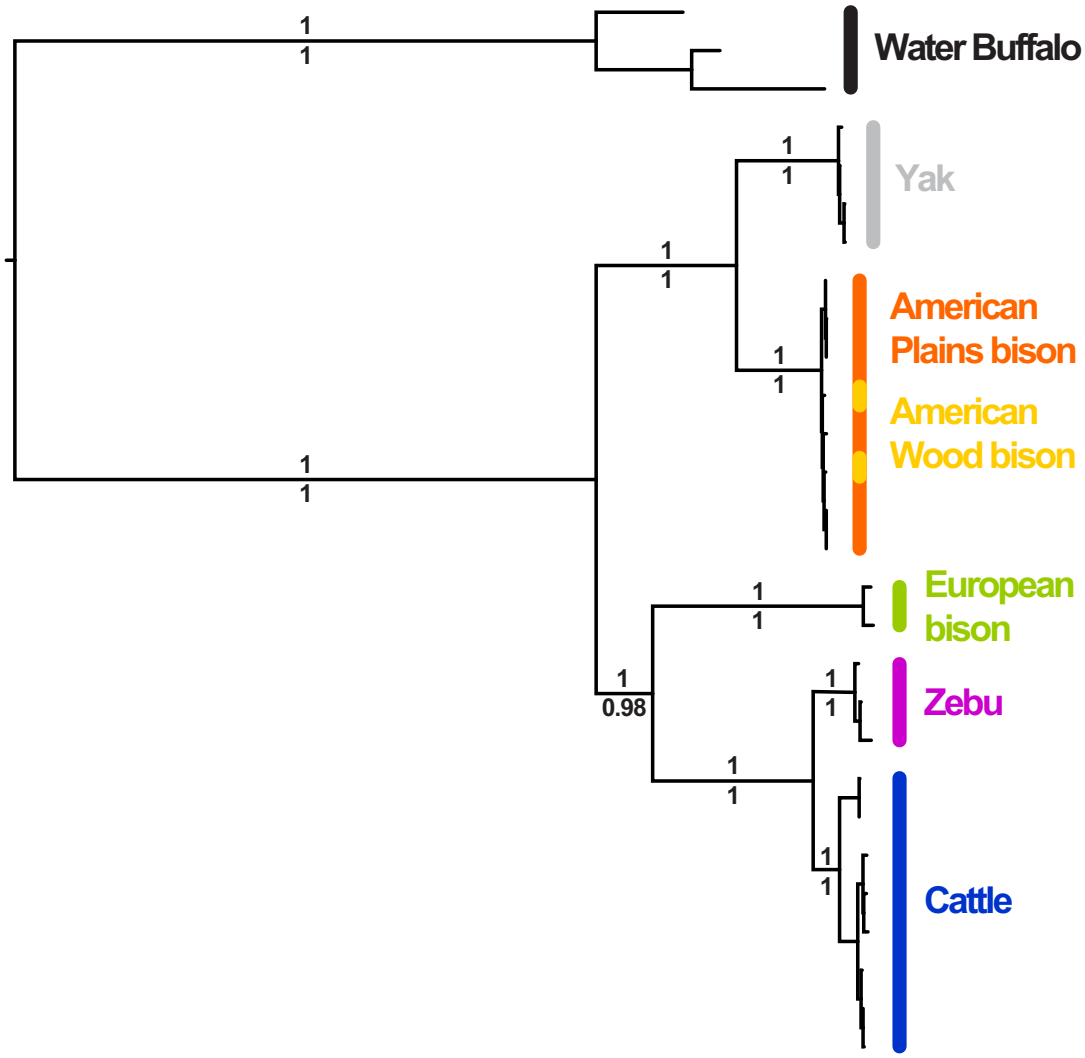


Figure 2. Bovid phylogeny derived from analysis of whole mitochondrial genome sequences, showing strong support for the grouping of European bison with cattle and zebu. Whole mitochondrial genome sequences are not currently available for BisonX. Numbers on top of branches represent the posterior probabilities from MrBayes, and numbers underneath represent aLRT support values from PhyML.

Two possible reasons have been proposed to explain the difference between the mitochondrial and nuclear topologies of the European bison: incomplete lineage sorting or hybridization event(s) (Verkaar et al. 2004). Incomplete mitochondrial lineage sorting is unlikely given the number of speciation events between European bison and cattle on the nuclear tree (gaur, banteng, yak, and American bison, see Figure 3 in Chapter 3). On the other hand, one (or multiple) introgression event(s) between the ancestral lineages of cattle/zebu and BisonX/European bison also seems somewhat unlikely. The hybridisation event would have had to be ancient to explain the branch lengths leading to European Bison and BisonX, and presumably would need to have occurred during a population bottleneck, or with repeated sex-biased (male *Bison*, female cattle) events, in order for the ancestral cattle mtDNA lineage to become fixed. However, modern cattle-European bison hybrids are not viable without medical intervention, and F1 males are infertile. Neither scenario currently seems likely.

Late Pleistocene movements of bison in Europe

The chronological distribution of bison specimens collected from the Urals, the North Sea and the Caucasus presented in Figure 3 revealed two periods of abrupt turnover and rapid replacement between Steppe bison and BisonX populations. Prior to the rapid temperature rise at Greenland Interstadial (GI)14 around 54 ky BP (dotted line on Figure 3), BisonX individuals are observed almost exclusively from all three locations. After 54 ky BP, Steppe bison specimens dominate the fossil record of the Urals and North Sea until around 32 ky BP, when they disappear permanently from the records of all three areas and only BisonX is observed. The transition date of 32 ky BP (second dotted line on Figure 3) corresponds to a second major climate change, at GI 5 and marks the beginning of an extended cold period, which is only very briefly interrupted by GI 4 and 3, and extends through the last glacial maximum (LGM). BisonX fossils are observed until the end of the LGM, *ca.* 14.5ky which is GI1, and then disappear from both the Urals and Caucasus record. The widespread temporal and geographic sampling suggests this pronounced pattern of alternating dominance between Steppe bison and BisonX populations is not a taphonomic artefact. Furthermore, the timing of the genetic transitions closely match marked climatic changes, with BisonX appearing in colder periods than Steppe bison. Although the record is far less complete in the North Sea and the Caucasus due to taphonomy and ancient DNA preservation, the temporal correlation appears quite remarkable over such a wide geographic region, from the western coast of Europe to the border of Asia and the Near East.

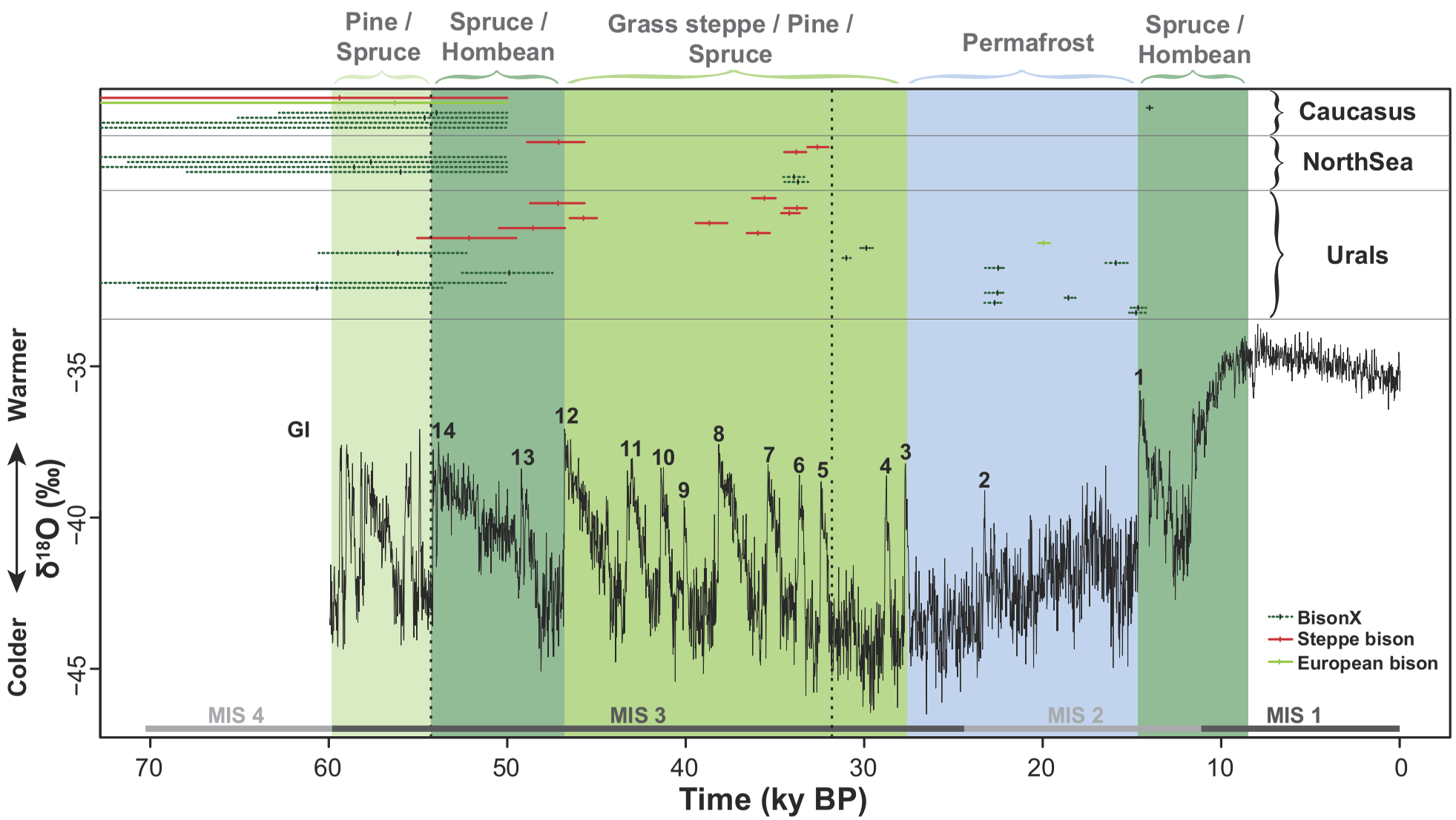

Pine / Spruce
Spruce / Hombean
Grass steppe / Pine / Spruce
Permafrost
Spruce / Hombean
Caucasus
NorthSea
Urals
Warmer
Colder
δ18O (‰)
−35
−40
−45
GI
14
13
12
11
10
9
8
7
6
5
4
3
2
1
BisonX
Steppe bison
European bison
MIS 4
MIS 3
MIS 2
MIS 1
70
60
50
40
30
20
10
0
Time (ky BP)

Figure 3. Geographical origin and chronology of the samples showing a series of replacement patterns that correlate with climate/paleovegetation events. Individual calibrated AMS dates for specimens (BisonX, Dark Green; Steppe bison *B. priscus* Red; European bison *B. bonasus* light green) are shown above for the three regions sampled (Urals, North Sea and Caucasus), with the NGRIP δO^{18} record below (Wolff et al. 2010). Greenland Interstadials (GI) are numbered in black, and Marine Isotope Stages (MIS) in grey. Paleovegetation reconstructions from Eifel, Germany (Sirocko et al. submitted) are superimposed, with descriptions at the top of the figure.
Towards the end of MIS Stage 4 (60 ky) BisonX is common across all three regions, but is then replaced by a dominance of Steppe Bison around GI14, ca. 55-48kyr, which is the warmest MIS3 interstadial in the paleovegetation reconstruction, and changed into a grass steppe with scattered pine and spruce which continued until GI3 (ca. 28kyr). There are no records from the Caucasus during this period. From ca. 32 ky (GI5) to 14.5 ky (GI1) Bison X is once again the dominant bovid observed in the Urals. These transitions closely match the change from steppe to LGM conditions – which were observed as a polar desert in the SW Germany site, and which may have occurred slightly earlier in the more northerly Urals. The last BisonX specimens (ca. 14.6 Urals – 14.0 Caucasus) are dated very close to the return to grass-steppe like conditions at Eifel after the LGM (ca. 14.5k). While Steppe bison occupied this environment in MIS 3, it was not detected after this stage and indeed was in severe population decline by GI1 (Shapiro et al. 2004). *Bison bonasus* (light green) was observed only twice, once in the MIS 4 layers at Mezmaiskaya and in the middle of the LGM in the Urals suggesting it was not common in Europe during MIS 4-2, and only became common in the Holocene.

The reconstructed paleovegetation history of Eifel, SW Germany covers the complete timeframe of this study and provides an interesting comparison although it is clearly distant from the Urals and Caucasus. The Eifel region records the presence of a classic steppe flora (grass steppe with scattered birch, pine and spruce) when the Steppe bison is present in the adjacent North Sea, as well as the Urals. In contrast, Bison X is present during colder periods characterised by either pine and spruce forest (earliest MIS 3) or a permafrost desert (MIS 2) in the Eifel area. The slight delay observed between the apparent transition between Steppe bison and Bison X populations at 32 ky in the Urals and North Sea, and the change from steppe vegetation to permafrost conditions around 28 ky at Eifel could well reflect regional differences, particularly in the higher latitudes of the Urals. Alternatively, the genetic transition occurs around GI5, at the beginning of an extended cold period that is only very briefly interrupted by GI 4 and 3 and continues through the LGM. As such, the bison populations may be responding to the temperature shifts more rapidly than the shifts in vegetation patterns at Eifel.

When the BisonX populations reappear at 32 ky BP, after an apparent 20 ky period of absence, they form a diverse but monophyletic group indicating a recent common genetic origin. This is consistent with the population surviving in a single refugium prior to this point, potentially in response to unfavourable environmental conditions. To investigate this scenario further, a temporally calibrated phylogeny of BisonX samples was constructed (Figure 4). The phylogeny was calibrated using the radiocarbon dates associated with the ancient sequences, after checking that the dates in question contained sufficient signal with a date randomization test (Supplementary Figure 1). The resulting phylogeny shows that the genetic diversity of BisonX at the end of MIS 4 (60 ky BP) could be divided into two clades, with the North Sea individuals forming a basal group, suggesting some degree of phylogeographic structure. However, following the warmer part of MIS 3 (between 53 and 32 ky BP), BisonX samples re-appear in the record with a comparable genetic diversity but no clear phylogeographic patterns. In fact, the initial reappearance of BisonX is recorded in the North Sea with two samples around 34 ky (see Figure 3 and 4), and these sequences already represent two extremes of the MIS 3 genetic diversity. This might suggest a refugium closer to the North Sea than the other sampled geographical locations, or that colder conditions re-appeared in the North Sea before the Urals. Bison X is then well represented in the Urals until 14.6 ky BP and appears in the Caucasus at 14 ky BP immediately before GI1 and the subsequent Younger Dryas. The time of the most recent common ancestor (tMRCA) of the MIS 3 BisonX diversity is estimated to be around 58 (45.8 – 74.9) ky (greyed bar on the tree). This date closely matches the ages of the last observed MIS 4 BisonX individuals across all

three locations, supporting the idea of a population movement of cold-adapted BisonX toward a refugium during the warmer period of MIS 3 in Europe.

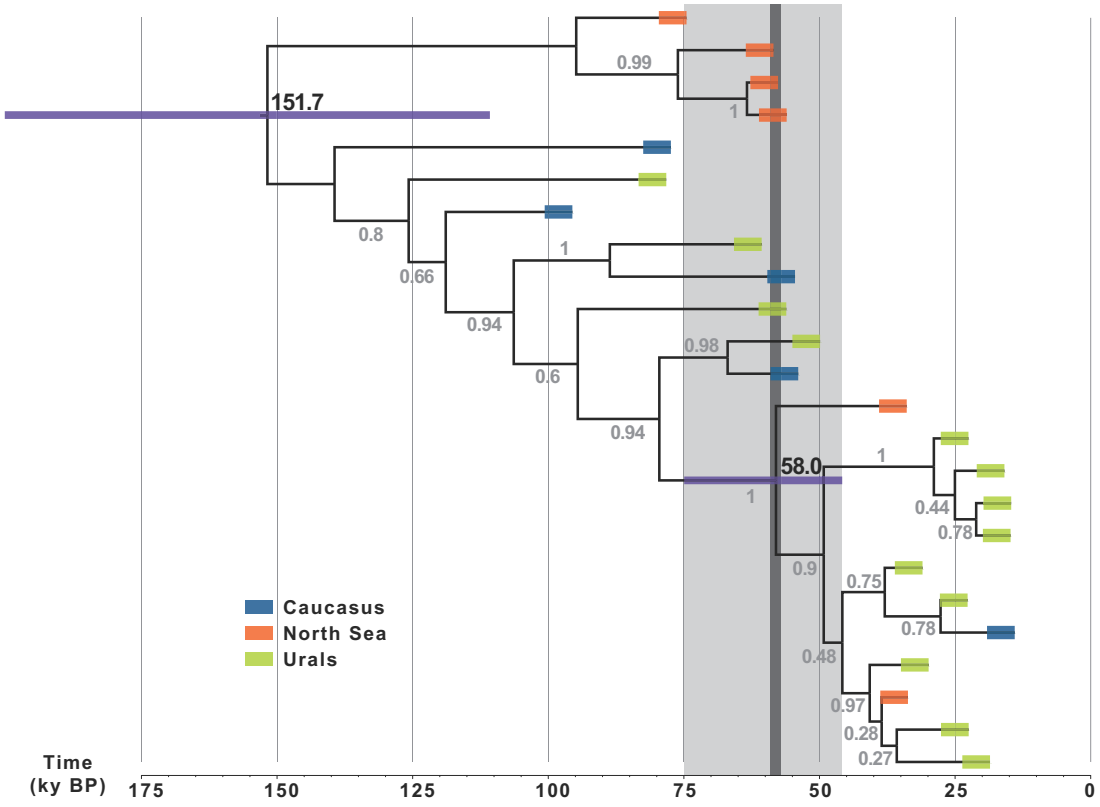


Figure 4. Maximum clade credibility tree of BisonX estimated using Bayesian analysis and calibrated with AMS radiocarbon dates associated with the sequenced bones. Dates older than 50kyr are inferred in the phylogenetic reconstruction, as specimens either returned infinite dates, or were not dated due to their stratigraphic position. Interestingly, the specimens which reappear in the colder conditions of late MIS 3/early MIS 2 (right of the grey highlighted region) form a derived clade within the tree indicating they share a recent common origin. Interestingly, the most recent specimens within the diversity observed during MIS 4 are dated around 55-60 kyr, close to the MIS 3-4 boundary, and the tMRCA of the MIS 3 BisonX clade (58 ky BP). This is consistent with a scenario in which BisonX retreated to a refugium during the warmer phases of MIS 3, and only re-expanded with the return of cold conditions during the earliest phases of the LGM.

While the Steppe bison is clearly associated with the warmer periods of MIS 3 after 54 ky BP, it is interesting that none of the 44 individuals detected in the Urals and North Sea are dated between 46 and 38 ky. This gap in the European bison record occurs just after a particularly cold stadial corresponding to Heinrich event 5a (50-48 ky BP). This particular time point of 48 ky has already been correlated with major changes in genetic diversity loss for several diverse Eurasian mammals, including Cave lions (Barnett et al. 2009), Neandertals (Dalén et al. 2012) and mammoths(Barnes et al. 2007; Gilbert et al. 2008). The paleovegetation record from Germany (Figure 3) and from a separate study in Greece (Müller et al. 2011) also show important alterations of the flora close to this date. The synchronicity of drastic perturbations in the temperature record, the vegetation record and the genetic diversity of large mammals (including hominids), suggests that important climate change remodelled the megafaunal diversity of Eurasia around 50 ky, well prior to the arrival of anatomically modern humans or the late Pleistocene extinctions following the LGM (Barnosky et al. 2004; Lorenzen et al. 2011).

While it is not currently possible to assign a firm taxonomic identity to BisonX, it is important to note that morphological studies in Dmanisi, Georgia, Voigtstedt and Ungermassfeld (Sher 1997; van der Made 1999). Germany, have identified a new early/mid Pleistocene bison species intermediate in size between *B. priscus* and *B. schoetensacki*. Consequently, both this species, termed *Bison cf. voigtstedtensis,* and *B. schoetensacki* appear potential candidates for the identity of BisonX.

References

Barnes, I, B Shapiro, A Lister, T Kuznetsova, A Sher, D Guthrie, MG Thomas. 2007. Genetic structure and extinction of the woolly mammoth, Mammuthus primigenius. Current biology : CB 17:1072-1075.

Barnett, R, B Shapiro, I Barnes, et al. 2009. Phylogeography of lions (Panthera leo ssp.) reveals three distinct taxa and a late Pleistocene reduction in genetic diversity. Molecular ecology 18:1668-1677.

Barnosky, AD, PL Koch, RS Feranec, SL Wing, AB Shabel. 2004. Assessing the causes of late Pleistocene extinctions on the continents. Science 306:70-75.

Barnosky, AD, EL Lindsey. 2010. Timing of Quaternary megafaunal extinction in South America in relation to human arrival and climate change. Quaternary International 217:10-29.

Barnosky, AD, N Matzke, S Tomiya, et al. 2011. Has the Earth's sixth mass extinction already arrived? Nature 471:51-57.

Bauer, K. 2001. Wisent Bison bonasus (Linnaeus, 1758). Graz, Austria: Bundesministerium für Land- und Forstwirtschaft Umwelt und Wasserwirtschaft.

Beja-Pereira, A, D Caramelli, C Lalueza-Fox, et al. 2006. The origin of European cattle: evidence from modern and ancient DNA. Proceedings of the National Academy of Sciences of the United States of America 103:8113-8118.

Bollongino, R, J Burger, A Powell, M Mashkour, J-D Vigne, MG Thomas. 2012. Modern Taurine Cattle descended from small number of Near-Eastern founders. Molecular Biology and Evolution.

Campos, PF, T Kristensen, L Orlando, et al. 2010. Ancient DNA sequences point to a large loss of mitochondrial genetic diversity in the saiga antelope (Saiga tatarica) since the Pleistocene. Molecular ecology 19:4863-4875.

Dalén, L, L Orlando, B Shapiro, MBm Durling, R Quam, MTP Gilbert, JC Díez Fernández-Lomana, E Willerslev, JL Arsuaga, A Götherström. 2012. Partial genetic turnover in neandertals: continuity in the east and population replacement in the west. Molecular Biology and Evolution.

Drummond, AJ, A Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC evolutionary biology 7:214.

Edwards, CJ, R Bollongino, A Scheu, et al. 2007. Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. Proceedings of the Royal Society B: Biological Sciences 274:1377-1385.

Edwards, CJ, DA Magee, SD Park, et al. 2010. A complete mitochondrial genome sequence from a mesolithic wild aurochs (Bos primigenius). Plos One 5:e9255.

Flerov, CC. 1979. European Bison. Morphology, Systematics, Evolution, Ecology. Moscow: Nauka Publishers [in Russian].

Geist, V, P Karsten. 1977. The wood bison in relation to hypothesis on the origin of the American bison. Z. Säugetierk 42:119-122.

Gilbert, MT, DI Drautz, AM Lesk, et al. 2008. Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes. Proceedings of the National Academy of Sciences of the United States of America 105:8327-8332.

Graham, RW, EL Lundelius, Jr., MA Graham, et al. 1996. Spatial Response of Mammals to Late Quaternary Environmental Fluctuations. Science 272:1601-1606.

Guindon, S, JF Dufayard, V Lefort, M Anisimova, W Hordijk, O Gascuel. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic biology 59:307-321.

Ho, SYW, R Lanfear, MJ Phillips, I Barnes, JA Thomas, SO Kolokotronis, B Shapiro. 2011. Bayesian estimation of substitution rates from ancient DNA sequences with low information content. Systematic biology 60:366-375.

Janecek, LL, RL Honeycutt, RM Adkins, SK Davis. 1996. Mitochondrial Gene Sequences and the Molecular Systematics of the Artiodactyl Subfamily Bovinae. Molecular phylogenetics and evolution 6:107-119.

Kahlke, R-D. 1999. The history of the origin, evolution and dispersal of the late Pleistocene Mammuthus-Coelodonta faunal complex in Eurasia (large mammals). Rapid City, South Dakota: Fenske Companies, Mammoth Site of Hot Springs, South Dakota, Inc.

Keane, TM, CJ Creevey, MM Pentony, TJ Naughton, JO McLnerney. 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. BMC evolutionary biology 6:29.

Kurtén, B. 1968. Pleistocene Mammals of Europe. London, England: Weidenfeld and Nicolson.

Lari, M, E Rizzi, S Mona, et al. 2011. The complete mitochondrial genome of an 11,450-year-old aurochsen (Bos primigenius) from Central Italy. BMC evolutionary biology 11:32.

Lazarev, PA, GG Boeskorov, AI Tomskaya. 1998. Mlekopitauschie Antropogena Yakutii. Yakutsk, Yakutskii nauchnyi tsentr: Yu.V.Labutin (Editor) Sibirskoe otdelenie Rossiiskoi Akademii Nauk.

Lorenzen, ED, D Nogues-Bravo, L Orlando, et al. 2011. Species-specific responses of Late Quaternary megafauna to climate and humans. Nature 479:359-364.

Martin, PS. 1984. Prehistoric overkill: the global model. Quaternary extinctions:354-403.

McDonald, JN. 1980. North American bison. Their classification and evolution. Berkeley: University of California Press.

Minin, VN, EW Bloomquist, MA Suchard. 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Molecular Biology and Evolution 25:1459-1471.

Müller, UC, J Pross, PC Tzedakis, C Gamble, U Kotthoff, G Schmiedl, S Wulf, K Christanis. 2011. The role of climate in the spread of modern humans into Europe. Quaternary Science Reviews 30:273-279.

Nystrom, V, L Dalen, S Vartanyan, K Liden, N Ryman, A Angerbjorn. 2010. Temporal genetic change in the last remaining population of woolly mammoth. Proceedings of the Royal Society B: Biological Sciences 277:2331-2337.

Orlando, L, D Bonjean, H Bocherens, A Thenot, A Argant, M Otte, C Hanni. 2002. Ancient DNA and the population genetics of cave bears (Ursus spelaeus) through space and time. Molecular Biology and Evolution 19:1920-1933.

Pucek, Z. 1986. Bison bonasus (Linnaeus 1758) - Wisent. In: FK J Niethammer, editor. Handbuch der Säugetiere Europas. Wiesbaden: Aula Verlag.

Reimer, PJ, MGL Baillie, E Bard, et al. 2009. IntCal09 and Marine09 radiocarbon age calibration curves, 0-50,000 years CAL BP. Radiocarbon 51:1111-1150.

Rohland, N, JL Pollack, D Nagel, C Beauval, J Airvaux, S Paabo, M Hofreiter. 2005. The population history of extant and extinct hyenas. Molecular Biology and Evolution 22:2435-2443.

Ronquist, F, M Teslenko, P van der Mark, DL Ayres, A Darling, S Hohna, B Larget, L Liu, MA Suchard, JP Huelsenbeck. 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. Systematic biology.

Rule, S, BW Brook, SG Haberle, CS Turney, AP Kershaw, CN Johnson. 2012. The aftermath of megafaunal extinction: ecosystem transformation in Pleistocene Australia. Science 335:1483-1486.

Shapiro, B, AJ Drummond, A Rambaut, et al. 2004. Rise and fall of the Beringian steppe bison. Science 306:1561-1565.

Shapiro, B, SYW Ho, AJ Drummond, MA Suchard, OG Pybus, A Rambaut. 2011. A Bayesian phylogenetic method to estimate unknown sequence ages. Molecular Biology and Evolution 28:879-887.

Sher, AV. 1997. An Early Quaternary bison population from Untermaßfeld: Bison menneri sp. nov. In: R-D Kahlke, editor. Das Pleistozän von Untermaßfeld bei Meiningen (Thüringen). Bonn: Habelt-Verlag. p. 101–180.

Skinner, AR, BA Blackwell, S Martin, A Ortega, JI Blickstein, LV Golovanova, VB Doronichev. 2005. ESR dating at Mezmaiskaya Cave, Russia. Applied radiation and isotopes : including data, instrumentation and methods for use in agriculture, industry and medicine 62:219-224.

Slatis, HM. 1960. An Analysis of Inbreeding in the European Bison. Genetics 45:275–287.

Stewart, JR, CB Stringer. 2012. Human evolution out of Africa: the role of refugia and climate change. Science 335:1317-1321.

Stiller, M, G Baryshnikov, H Bocherens, et al. 2010. Withering away--25,000 years of genetic decline preceded cave bear extinction. Molecular Biology and Evolution 27:975-978.

Stuart, AJ. 1991. Mammalian extinctions in the late Pleisticene of northern Eurasia and North America. Biological Reviews 66:453-562.

Stuart, AJ, AM Lister. 2007. Patterns of Late Quaternary megafaunal extinctions in Europe and northern Asia. In: RD Kahlke, LC Maul, PPA Mazza, editors. Late Neogene and Quaternary Biodiversity and Evolution: Regional Developments and Interregional Correlations, Vol Ii. Frankfurt: Senckenbergische Naturforschende Gesellschaft. p. 287-297.

van der Made, J. 1999. Ungulates from Atapuerca TD6. Journal of Human Evolution 37:389-413.

Verkaar, EL, IJ Nijman, M Beeke, E Hanekamp, JA Lenstra. 2004. Maternal and paternal lineages in cross-breeding bovine species. Has wisent a hybrid origin? Molecular Biology and Evolution 21:1165-1170.

Wolff, EW, J Chappellaz, T Blunier, SO Rasmussen, A Svensson. 2010. Millennial-scale variability during the last glacial: The ice core record. Quaternary Science Reviews 29:2828-2838.

Acknowledgments:

The authors would like to thank Pavel Kosintsev, Alex Vorobiev, Jan Glimmerveen, Liubov Golovanova and Vladimir Doronichev for providing samples, and Laura Watson and Gaynor Dolman for their help in ancient DNA work.

Supplementary material

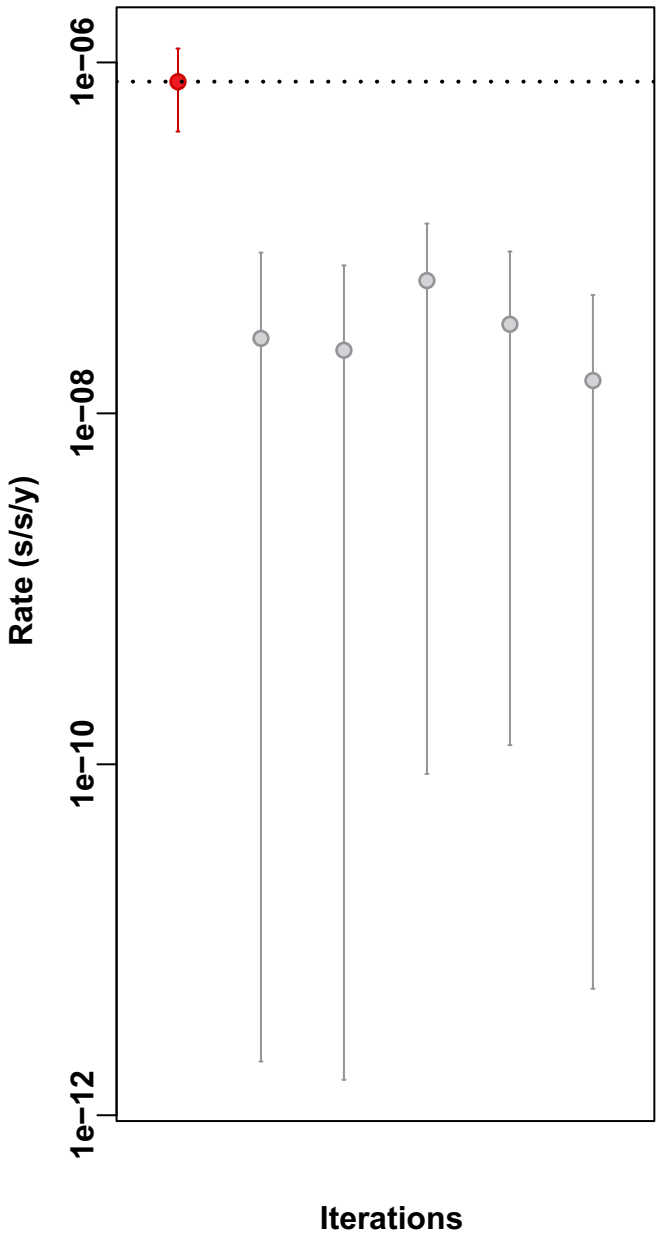


Figure S1. Date randomization test. The red circle and dotted line represent the molecular rate estimated during the phylogenetic analysis of BisonX using the radiocarbon dates associated with the ancient sequences as calibration. The grey lines represent the 95% HPD of rates estimated with randomized dates. The fact that none of these margins overlap with the original mean rate demonstrates that the radiocarbon dates used for this study is informative enough to calibrate the tree.

Table S1. List of published mitochondrial control region sequences used for phylogenetic analysis. The Urals Steppe bison are highlighted in red. The two last numbers correspond to sample date and the associated error margin.

American bison	Bison_priscus_BS146_NS_11810_50	Bison_priscus_BS397_NS_32370_470	Bos_indicus_AY378135_0_0
Bison_bison_AF083357_H1_0_0	Bison_priscus_BS147_NS_28120_290	Bison_priscus_BS398_NS_27400_260	Bos_indicus_DQ887765_0_0
Bison_bison_AF083358_H2_0_0	Bison_priscus_BS148_NS_6400_50	Bison_priscus_BS400_NS_46100_2600	Bos_indicus_EF417971_0_0
Bison_bison_AF083359_H3_0_0	Bison_priscus_BS149_NS_46100_2200	Bison_priscus_BS405_SI_23040_120	Bos_indicus_EF417974_0_0
Bison_bison_AF083360_H4_0_0	Bison_priscus_BS150_NS_10510_50	Bison_priscus_BS407_NWT_55500_3100	Bos_indicus_EF417976_0_0
Bison_bison_AF083361_H5_0_0	Bison_priscus_BS151_NS_21530_130	Bison_priscus_BS412_Y_30500_250	Bos_indicus_EF417977_0_0
Bison_bison_AF083362_H6_0_0	Bison_priscus_BS161_NS_21040_120	Bison_priscus_BS414_BIR_4495_60	Bos_indicus_EF417979_0_0
Bison_bison_AF083363_H7_0_0	Bison_priscus_BS163_LC_13240_75	Bison_priscus_BS415_D_30810_975	Bos_indicus_EF417981_0_0
Bison_bison_AF083364_H8_0_0	Bison_priscus_BS164_LC_19540_120	Bison_priscus_BS418_China_26560_670	Bos_indicus_EF417983_0_0
Bison_bison_BS100_29_5	Bison_priscus_BS165_LC_26460_160	Bison_priscus_BS438_AB_53800_2200	Bos_indicus_EF417985_0_0
Bison_bison_BS102_22_5	Bison_priscus_BS170_YT_13040_70	Bison_priscus_BS440_AB_60400_2900	Bos_indicus_EF524120_0_0
Bison_bison_BS129_0_2000	Bison_priscus_BS172_LC_12525_70	Bison_priscus_BS443_AB_34050_450	Bos_indicus_EF524125_0_0
Bison_bison_BS162_AK_170_30	Bison_priscus_BS176_LC_12380_60	Bison_priscus_BS459_China_47700_1000	Bos_indicus_EF524126_0_0
Bison_bison_BS173_NTC_3220_45	Bison_priscus_BS178_LC_17960_90	Bison_priscus_BS469_AB_305_24	Bos_indicus_EF524128_0_0
Bison_bison_BS175_ICE_186_30	Bison_priscus_BS192_F_26300_300	Bison_priscus_BS472_F_13235_65	Bos_indicus_EF524130_0_0
Bison_bison_BS177_NTC_3155_36	Bison_priscus_BS193_NS_49600_4000	Bison_priscus_BS473_AB_56300_3100	Bos_indicus_EF524132_0_0
Bison_bison_BS200_AB_145_37	Bison_priscus_BS195_NS_29040_340	Bison_priscus_BS477_D_33710_240	Bos_indicus_EF524135_0_0
Bison_bison_BS342_CHL_10340_40	Bison_priscus_BS196_NS_19420_100	Bison_priscus_BS478_D_34470_200	Bos_indicus_EF524141_0_0
Bison_bison_BS348_CHL_10505_45	Bison_priscus_BS198_Y_2460_40	Bison_priscus_BS490_BIR_2415_25	Bos_indicus_EF524152_0_0
Bison_bison_BS368_0_2000	Bison_priscus_BS201_Y_12960_60	Bison_priscus_BS493_NS_50000_4200	Bos_indicus_EF524156_0_0
Bison_bison_BS417_AB_909_29	Bison_priscus_BS202_AB_10460_65	Bison_priscus_BS494_NS_44800_2200	Bos_indicus_EF524160_0_0
Bison_bison_BS419_AB_7475_45	Bison_priscus_BS206_Sibh_23780_140	Bison_priscus_BS495_NS_29570_340	Bos_indicus_EF524166_0_0
Bison_bison_BS421_AB_8145_45	Bison_priscus_BS211_Sibh_43800_1100	Bison_priscus_BS497_NS_30000_540	Bos_indicus_EF524167_0_0
Bison_bison_BS422_AB_908_31	Bison_priscus_BS216_NS_47000_2900	Bison_priscus_BS498_NS_25980_230	Bos_indicus_EF524170_0_0
Bison_bison_BS423_AB_4660_38	Bison_priscus_BS218_SI_14605_75	Bison_priscus_BS499_NS_31410_420	Bos_indicus_EF524177_0_0
Bison_bison_BS424_AB_202_32	Bison_priscus_BS222_NWT_6110_45	Bison_priscus_BS500_NS_35580_550	Bos_indicus_EF524180_0_0
Bison_bison_BS426_AB_7060_45	Bison_priscus_BS223_SI_53300_1900	Bison_priscus_BS517_BIR_2526_26	Bos_indicus_EF524183_0_0
Bison_bison_BS428_AB_7105_45	Bison_priscus_BS224_AK_13125_75	Bison_priscus_BS564_SI_24570_90	Bos_indicus_EF524185_0_0
Bison_bison_BS429_AB_6775_40	Bison_priscus_BS233_SW_16685_80	Bison_priscus_BS571_SIdy_32910_170	Bos_indicus_L27732_0_0
Bison_bison_BS430_9270_50	Bison_priscus_BS235_BIR_43400_900	Bison_priscus_BS592_Urals_42500_450	Bos_indicus_L27736_0_0
Bison_bison_BS432_AB_7310_45	Bison_priscus_BS236_SW_19420_100	Bison_priscus_BS605_NTC_20380_90	**Aurochs**
Bison_bison_BS433_AB_10450_55	Bison_priscus_BS237_AB_11240_70	Bison_priscus_BS660_Urals_29500_140	Bos_primigenius_DQ915522_ALL1_12030_52
Bison_bison_BS434_AB_809_32	Bison_priscus_BS243_SW_37550_400	Bison_priscus_BS662_SI_20000_0	Bos_primigenius_DQ915523_CAT1_5650_0
Bison_bison_BS439_AB_5845_45	Bison_priscus_BS244_LC_26210_170	Bison_priscus_BS674_Urals_29060_140	Bos_primigenius_DQ915524_CHWF_3905_185
Bison_bison_BS441_AB_1273_32	Bison_priscus_BS248_OCr_12350_70	Bison_priscus_BS708_Urals_47050_750	Bos_primigenius_DQ915537_CPC98_5936_34
Bison_bison_BS444_AB_636_29	Bison_priscus_BS249_F_39200_550	Bison_priscus_BS713_Urals_30970_180	Bos_primigenius_DQ915542_EIL06_5830_29
Bison_bison_BS445_AB_378_30	Bison_priscus_BS253_LC_12665_65	Bison_priscus_IB179_LC_12465_75	Bos_primigenius_DQ915543_EIL14_5830_29
Bison_bison_BS449_6195_45	Bison_priscus_BS254_CHL_10230_55	**European bison**	Bos_primigenius_DQ915554_LJU3_8020_50
Bison_bison_BS454_AB_287_29	Bison_priscus_BS258_F_22120_130	Bison_bonasus_AF083356_0_0	Bos_primigenius_DQ915558_NORF_3370_30
Bison_bison_BS456_AB_125_30	Bison_priscus_BS260_D_30750_290	Bison_bonasus_AY428860_0_0	Bos_primigenius_EF187280_PVL04_3204_56
Bison_bison_BS460_AB_10425_50	Bison_priscus_BS261_LC_12915_70	Bison_bonasus_EF693811_0_0	**Cattle**
Bison_bison_BS464_AB_5205_45	Bison_priscus_BS262_D_29150_500	Bison_bonasus_EU272053_0_0	Bos_taurus_DQ124372_T4_0_0
Bison_bison_BS465_AB_7115_50	Bison_priscus_BS281_BIR_40800_600	Bison_bonasus_EU272054_0_0	Bos_taurus_DQ124375_T4_0_0
Bison_bison_BS466_AB_3298_37	Bison_priscus_BS282_SI_56700_3200	Bison_bonasus_EU272055_0_0	Bos_taurus_DQ124381_T3_0_0
Bison_bison_BS503_BIR_2776_36	Bison_priscus_BS284_Y_13135_65	Bison_bonasus_U12953_0_0	Bos_taurus_DQ124383_T2_0_0
Bison_bison_BS560_AB_2807_28	Bison_priscus_BS286_Sim_49500_1300	Bison_bonasus_U12954_0_0	Bos_taurus_DQ124388_T3_0_0
Bison_bison_BS569_AB_3600_70	Bison_priscus_BS287_BIR_49100_1700	Bison_bonasus_U34294_0_0	Bos_taurus_DQ124394_T3_0_0
Bison_bison_BS570_AB_11300_290	Bison_priscus_BS289_BIR_2172_37	**Yak**	Bos_taurus_DQ124398_T3_0_0
Bison_bison_BS99_26_5	Bison_priscus_BS291_NS_49700_1400	Bos_grunniens_AY521140_0_0	Bos_taurus_DQ124400_T4_0_0
Bison_bison_U12935_0_0	Bison_priscus_BS292_NS_35710_730	Bos_grunniens_AY521149_0_0	Bos_taurus_DQ124401_T4_0_0
Bison_bison_U12936_0_0	Bison_priscus_BS294_BIR_58200_3900	Bos_grunniens_AY521150_0_0	Bos_taurus_DQ124412_T4_0_0
Bison_bison_U12941_0_0	Bison_priscus_BS297_NS_10990_50	Bos_grunniens_AY521151_0_0	Bos_taurus_EU177822_T3_0_0
Bison_bison_U12943_0_0	Bison_priscus_BS311_BIR_12425_45	Bos_grunniens_AY521152_0_0	Bos_taurus_EU177841_T1_0_0
Bison_bison_U12944_0_0	Bison_priscus_BS316_SI_57700_3000	Bos_grunniens_AY521154_0_0	Bos_taurus_EU177842_T1_0_0
Bison_bison_U12945_0_0	Bison_priscus_BS318_NS_12410_50	Bos_grunniens_AY521155_0_0	Bos_taurus_EU177845_T1_0_0
Bison_bison_U12946_0_0	Bison_priscus_BS320_SI_49600_1500	Bos_grunniens_AY521156_0_0	Bos_taurus_EU177847_T1_0_0
Bison_bison_U12947_0_0	Bison_priscus_BS321_AK_9506_38	Bos_grunniens_AY521160_0_0	Bos_taurus_EU177848_T1_0_0
Bison_bison_U12948_0_0	Bison_priscus_BS323_SI_37810_380	Bos_grunniens_AY521161_0_0	Bos_taurus_EU177853_T2_0_0
Bison_bison_U12955_0_0	Bison_priscus_BS327_D_31530_230	Bos_grunniens_DQ007210_0_0	Bos_taurus_EU177854_T2_0_0
Bison_bison_U12956_0_0	Bison_priscus_BS328_SIdy_31690_180	Bos_grunniens_DQ007221_0_0	Bos_taurus_EU177860_T2_0_0
Bison_bison_U12957_0_0	Bison_priscus_BS329_D_27060_190	Bos_grunniens_DQ007222_0_0	Bos_taurus_EU177861_T2_0_0
Bison_bison_U12958_0_0	Bison_priscus_BS337_CHL_10378_36	Bos_grunniens_DQ856594_0_0	Bos_taurus_EU177862_T5_0_0
Bison_bison_U12959_0_0	Bison_priscus_BS340_NS_24500_180	Bos_grunniens_DQ856599_0_0	Bos_taurus_EU177863_T5_0_0
Steppe bison	Bison_priscus_BS345_NS_39800_1200	Bos_grunniens_DQ856600_0_0	Bos_taurus_EU177864_T5_0_0
Bison_priscus_A3133_Yukon_26360_220	Bison_priscus_BS350_NS_38700_1000	Bos_grunniens_DQ856603_0_0	Bos_taurus_EU177865_T5_0_0
Bison_priscus_BS105_F_23380_460	Bison_priscus_BS351_BIR_57700_3200	Bos_grunniens_DQ856604_0_0	**Buffalo**
Bison_priscus_BS107_F_19570_290	Bison_priscus_BS359_NTC_20020_150	Bos_grunniens_EF494177_0_0	Bubalus_bubalis_AF197208_0_0
Bison_priscus_BS108_F_21020_360	Bison_priscus_BS364_NS_38800_1100	Bos_grunniens_EF494178_0_0	Bubalus_bubalis_AF475212_0_0
Bison_priscus_BS109_F_20730_350	Bison_priscus_BS365_NS_47000_2900	**Zebu**	Bubalus_bubalis_AF475256_0_0
Bison_priscus_BS111_F_21580_370	Bison_priscus_BS387_NS_33320_540	Bos_indicus_AB085923_0_0	Bubalus_bubalis_AF475259_0_0
Bison_priscus_BS121_F_19360_280	Bison_priscus_BS388_NS_27590_280	Bos_indicus_AB268563_0_0	Bubalus_bubalis_AF475278_0_0
Bison_priscus_BS123_BIR_1730_60	Bison_priscus_BS389_NS_17160_80	Bos_indicus_AB268564_0_0	Bubalus_bubalis_AY488491_0_0
Bison_priscus_BS124_BIR_11900_70	Bison_priscus_BS390_NS_31630_440	Bos_indicus_AB268566_0_0	Bubalus_bubalis_EF536327_0_0
Bison_priscus_BS125_F_27440_790	Bison_priscus_BS392_NS_36320_780	Bos_indicus_AB268571_0_0	Bubalus_bubalis_EF536328_0_0
Bison_priscus_BS126_F_19150_280	Bison_priscus_BS393_NS_39850_1200	Bos_indicus_AB268574_0_0	Bubalus_bubalis_EU268899_0_0
Bison_priscus_BS130_BIR_9000_250	Bison_priscus_BS394_NS_37460_890	Bos_indicus_AB268578_0_0	Bubalus_bubalis_EU268909_0_0
Bison_priscus_BS133_F_33800_1900	Bison_priscus_BS395_NS_40700_1300	Bos_indicus_AB268580_0_0	
Bison_priscus_BS145_NS_12270_50	Bison_priscus_BS396_NS_23680_170	Bos_indicus_AY378134_0_0	

Table S2. List of published whole mitochondrial genome sequences used for phylogenetic analysis.

American bison	Cattle	Buffalo	Yak
GU947000_Bison_bison_PlainNebraska	FJ971081_Bos_taurus_Chianina	AY488491_Bubalus_bubalis	EF494178_Bos_grunniens
GU946976_Bison_bison_PlainMontana	FJ971083_Bos_taurus_Romagnola	AY702618_Bubalus_bubalis	EF494177_Bos_grunniens
GU947004_Bison_bison_PlainWyoming	FJ971082_Bos_taurus_ItalianRedPied	AF547270_Bubalus_bubalis	EF494179_Bos_grunniens
GU947006_Bison_bison_WoodElkIsland14	GU947021_Bos_taurus_Longhorn	**Zebu**	GQ464310_Bos_grunniens
GU946987_Bison_bison_PlainMontana	AY676864_Bos_taurus_Angus	GU256940_Bos_indicus	**European bison**
GU947005_Bison_bison_WoodElkIsland1	AB074965_Bos_taurus_JapaneseBlack	AY126697_Bos_indicus	HQ223450_Bison_bonasus
GU947002_Bison_bison_PlainTexas	FJ971085_Bos_taurus_Cinisara	AF492350_Bos_indicus	HM045017_Bison_bonasus
GU947003_Bison_bison_PlainTexas	FJ971084_Bos_taurus_Agerolese		

Table S3. Primers used in this study

	Primer	Primer Sequence (5' --- 3')	Length[(c)]
Set_A1	BovCR-16351F	CAACCCCCAAAGCTGAAG	~96bp
	BovCR-16457R	TGGTTRGGGTACAAAGTCTGTG	
Set_B1	BovCR-16420F	CCATAAATGCAAAGAGCCTCAYCAG	~172bp
	BovCR-16642R	TGCATGGGGCATATAATTTAATGTA	
Set_A2	BovCR-16507F	AATGCATTACCCAAACRGGG	~184bp
	BovCR-16755R	ATTAAGCTCGTGATCTARTGG	
Set-B2	BovCR-16633F[(a)]	GCCCCATGCATATAAGCAAG	~132bp
	BovCR-16810R[(a)]	GCCTAGCGGGTTGCTGGTTTCACGC	
Set_A3	BovCR-16765F[(a)]	GAGCTTAAYTACCATGCCG	~125bp
	BovCR-16998R	CGAGATGTCTTATTTAAGAGGAAAGAATGG	
Set_B3	BovCR-16960F	CATCTGGTTCTTTCTTCAGGGCC	~110bp
	BovCR-80R[(a)]	CAAGCATCCCCCAAAATAAA	

Two pairs of PCR primers derived from hypervariable control region and 12S-rRNA region of the mitochondrial genome were used for one-step simplex PCRs.

	Primer	Primer Sequence (5'--- 3')	Length [(c)]
Frag1	BovCR_16738MF[(b)]	***CACGACGTTGTAAAACGAC***ATYGTACATAGYA CATTATGTCAA	67bp
	BovCR_16810TR[(b)]	***TACGACTCACTATAGGGCGA***GCCTAGCGGGTT GCTGGTTTCACGC	
Frag2	Mamm_12SE	CTATAATCGATAAACCCCGATA	96bp
	Mamm_12SH	GCTACACCTTGACCTAAC	

(a): Primers (BovCR-16633F, BovCR-16810R, BovCR-16765F, BovCR-80R) were published in (Shapiro et al. 2004).

(b): To obtain good quality sequences for short fragment from directly sequencing, M13 (CAC GAC GTT GTA AAA CGA C) and T7 (TAC GAC TCA CTA TAG GGC GA) primer sequences were tagged at the primers BovCR_16738F and BovCR_16810R, respectively.

(c): Length of PCR amplicon is primer-excluded.

Chapter 5

Study of the time dependence of molecular rates

(Part 1)

Time-dependent rates of molecular evolution

Statement of authorship

Time-dependent rates of molecular evolution

Simon Y. W. Ho

Initiated and led the review project, wrote the manuscript, developed the figures and acted as the corresponding author.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… …… *Date* ……05.04.2012……

Robert Lanfear

Contributed ideas, helped developing figure and writing the manuscript.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed … *Date* 7/4/12

Lindell Bromham

Contributed ideas and helped to revise the manuscript.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… …… *Date* ..

Matthew J. Phillips

Contributed ideas and helped writing the manuscript.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ….. …… *Date* 09/04/12

Julien Soubrier (Candidate)

Contributed to the organization of ideas and manuscript writing.

I hereby certify that the statement of contribution is accurate.

Signed …… …… Date 02.04.12

Allen G. Rodrigo

Contributed ideas.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ……. …… Date ……6 April 2012……

Alan Cooper

Contributed ideas and helped to revise the manuscript.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .. …… *Date ……05.04.2012……*

MOLECULAR ECOLOGY

Molecular Ecology (2011) doi: 10.1111/j.1365-294X.2011.05178.x

INVITED REVIEW

Time-dependent rates of molecular evolution

SIMON Y. W. HO,*† ROBERT LANFEAR,* LINDELL BROMHAM,* MATTHEW J. PHILLIPS,*[1] JULIEN SOUBRIER,‡ ALLEN G. RODRIGO§¶ and ALAN COOPER‡

Centre for Macroevolution and Macroecology, Evolution Ecology & Genetics, Research School of Biology, Australian National University, Canberra, ACT, Australia, †School of Biological Sciences, Edgeworth David Building A11, University of Sydney, Sydney, NSW, Australia, ‡Australian Centre for Ancient DNA, School of Earth & Environmental Sciences, University of Adelaide, Adelaide, SA, Australia, §Department of Biology, Duke University, Durham, NC, USA, ¶Bioinformatics Institute and School of Biological Sciences, University of Auckland, Auckland, New Zealand

Abstract

For over half a century, it has been known that the rate of morphological evolution appears to vary with the time frame of measurement. Rates of microevolutionary change, measured between successive generations, were found to be far higher than rates of macroevolutionary change inferred from the fossil record. More recently, it has been suggested that rates of molecular evolution are also time dependent, with the estimated rate depending on the timescale of measurement. This followed surprising observations that estimates of mutation rates, obtained in studies of pedigrees and laboratory mutation-accumulation lines, exceeded long-term substitution rates by an order of magnitude or more. Although a range of studies have provided evidence for such a pattern, the hypothesis remains relatively contentious. Furthermore, there is ongoing discussion about the factors that can cause molecular rate estimates to be dependent on time. Here we present an overview of our current understanding of time-dependent rates. We provide a summary of the evidence for time-dependent rates in animals, bacteria and viruses. We review the various biological and methodological factors that can cause rates to be time dependent, including the effects of natural selection, calibration errors, model misspecification and other artefacts. We also describe the challenges in calibrating estimates of molecular rates, particularly on the intermediate timescales that are critical for an accurate characterization of time-dependent rates. This has important consequences for the use of molecular-clock methods to estimate timescales of recent evolutionary events.

Keywords: calibrations, divergence times, molecular clock, mutation rate, purifying selection, substitution rate

Received 23 February 2011; revision received 24 May 2011; accepted 1 June 2011

Introduction

More than half a century ago, Kurtén (1959) found an inverse correlation between the rate of morphological evolution and the time interval over which the rate was measured. Specifically, the rate of morphological change between successive generations exceeded macroevolutionary rates by several orders of magnitude. This time-dependent rate pattern was confirmed in a number of subsequent studies, which consistently showed that morphological evolutionary rates appeared faster when measured over shorter timescales (Gingerich 1983, 2001; Roopnarine 2003). A similar phenomenon has recently emerged in rates of molecular evolution estimated from DNA sequence data. In particular, there is a striking disparity between spontaneous mutation rates, measured over a small number of generations in studies of pedigrees and laboratory mutation-accumulation lines, and the much lower substitution rates measured over geological time frames (e.g., Parsons *et al.* 1997; Howell *et al.* 2003; Santos *et al.* 2005; Gibbs *et al.* 2009).

Correspondence: Simon Y. W. Ho, Fax: +61 2 91140979; E-mail: simon.ho@sydney.edu.au

The discrepancy between short- and long-term rates raises an obvious question: for how long does the short-term rate remain elevated? Following the initial studies of mutation rates in human pedigrees that were published in the late 1990s (e.g., Howell *et al.* 1996; Parsons *et al.* 1997), some authors suggested that high rates persist for only a few generations (Macaulay *et al.* 1997; Gibbons 1998). However, a growing number of empirical studies have found that rates remain elevated for prolonged periods, perhaps exceeding a hundred thousand generations in some taxa (e.g., Ho *et al.* 2005; Genner *et al.* 2007; Burridge *et al.* 2008; Papadopoulou *et al.* 2010). Some researchers have postulated that rates of evolution decline exponentially as a function of the age of the calibration used to estimate them (Fig. 1; Ho & Larson 2006; Ho *et al.* 2005; Penny 2005), with young calibrations yielding rate estimates reflecting (nonlethal) mutation rates and older calibrations giving rate estimates reflecting substitution rates.

A comprehensive characterization of time dependence remains elusive because it requires that molecular rates be quantified accurately across a range of timescales. In turn, this depends on the availability of reliable temporal information for calibrating rate estimates. While such calibrations can be found for analyses involving time frames that are either very short (e.g., documented pedigrees and laboratory mutation-accumulation lines) or very long (e.g., palaeontological evidence), there is a paucity of reliable calibrations for intermediate time frames (see Box 1). However, estimating rates on intermediate time frames is crucial for studying the decay of high short-term rates.

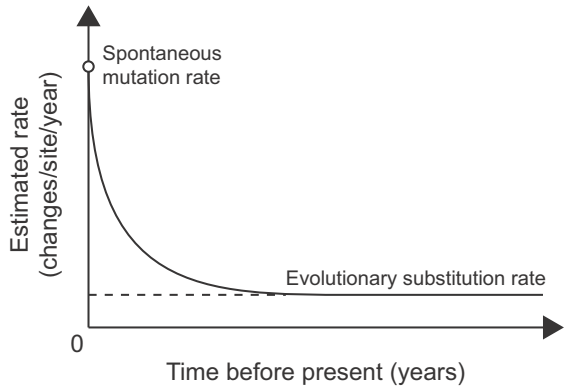


Fig. 1 Plot of time-dependent rates showing an exponentially declining rate estimate with increasing time depth. The spontaneous rate of non-lethal mutations is approached at a time depth of zero. As the time frame increases, the estimated rate tends towards the long-term substitution rate observed in phylogenetic analyses calibrated using palaeontological or geological data. The exact form of the curve is likely to show considerable variation among taxa and among loci.

Validating and quantifying the time dependence of molecular rates has important consequences for studies in a number of fields. One of the key implications is that rates estimated over a given time frame are not transferable to analyses of other time frames, meaning that it is invalid to assume that a single molecular clock applies throughout all lineages over time (e.g., Ho & Larson 2006; Howell *et al.* 2008). The degree of this problem is greatest for studies that use molecular dating to understand recent evolutionary events, for example those concerning viral phylodynamics, phylogeographic processes, prehistoric human dispersal or the extinction of megafauna in the late Pleistocene. The impact of time-dependent rates extends beyond the inference of timescales, however, and can also affect the inference of demographic parameters that use estimates of the mutation rate, such as effective population size and migration rates.

To be able to address the problem of time dependence when estimating molecular rates, timescales and other parameters, it is necessary to have a thorough appreciation of the causes of elevated short-term rates. There are many factors that can lead to elevated short-term rates, but their relative importance will vary among taxa. Disentangling these factors has wider implications for understanding the molecular evolutionary process (e.g., Woodhams 2006; Peterson & Masel 2009; O'Fallon *et al.* 2010).

Here we provide an overview of the known biological and methodological factors that can contribute to the time dependence of molecular evolutionary rates. We also list the various sources of information for calibrating rate estimates (Box 1) and survey the published evidence for time-dependent rates from studies of animals (Box 2) and bacteria and viruses (Box 3).

The basic biological framework

One of the most important biological reasons why we should expect molecular rate estimates to be time dependent is that rates measured on different timescales reflect different biological processes. Rates measured over very short timescales (e.g., between successive generations) can include genetic differences representing all but the most detrimental mutations. Therefore, these rate estimates approach the spontaneous mutation rate (discounting lethal mutations). On the other hand, rates measured over very long timescales (e.g., between distantly related species) will usually be dominated by substitutions (those mutations that were fixed in either diverging lineage) and will therefore approximate the substitution rate. Substitution rates are usually much lower than mutation rates because natural selection tends to remove delete-

Box 1. Calibrating estimates of molecular evolutionary rates

The degree of divergence between two sequences is determined by two factors: the rate of change and the time since they last shared a common ancestor. Therefore, to estimate the rate of molecular change, it is necessary to include independent information about the evolutionary timescale. Age calibrations can come from a variety of sources, including the fossil record, dated geological events, archaeological evidence, heterochronous sampling, documented pedigrees and laboratory lines (Figure 2).

Fossil record – One of the most common methods of calibration, particularly in studies of vertebrates, is the use of palaeontological evidence. Fossil evidence provides the earliest appearance of identifiable members of a lineage, allowing a minimum age constraint to be specified for the divergence event that gave rise to the lineage. If the fossil record is sufficiently informative, it is possible to include maximum age constraints or to give calibrations in the form of parametric age distributions (Yang & Rannala 2006; Ho & Phillips 2009).

In most cases, calibrations based on the fossil record are at least several million years in age. On shorter timescales, there is usually an insufficient amount of inherited morphological variation for the fossil record to provide reliable diagnostic characters. Consequently, palaeontological calibrations find their primary employment in phylogenetic analyses conducted above the population level.

Dated geological events – In a biogeographic context, divergence of species or populations can sometimes be attributed to geophysical isolating mechanisms or the appearance of new habitats. These can result from the formation of islands, mountain ranges, seaways or other geological features. If the timing of such an event is known, for example through radiometric dating, this information can be used to calibrate phylogenetic estimates of rates. Geological calibrations span a wide range of ages and can include island formation events and ancient continental movements.

Biogeographic calibrations can come in several forms, depending on the nature of the geological event and its impact on lineage divergences; for example, the appearance of the Panamanian isthmus might represent a barrier to gene flow between the Caribbean Sea and the Pacific Ocean (minimum age constraint), or a land bridge between the American continents (maximum age constraint). It is not always clear whether it is safe to assume a close correspondence between the age of the geological event and genetic divergence, because disparities can arise as a result of ancestral divergence or subsequent dispersal (Marko 2002; Heads 2005).

Archaeological and anthropological evidence – Molecular estimates of the timing of human migration have sometimes utilized calibrations informed by archaeological evidence. These calibrations include the arrival of humans in Australasia (about 45 000 years ago) and the Americas (about 14 000 years ago), as well as other migration events spanning a range of geographic and temporal scales (Endicott & Ho 2008; Henn *et al.* 2009; Subramanian 2009).

Diagnoses of human infection from archaeological sites or from historical documentation have been used to provide calibrations for analyses of viral and bacterial evolution; for example, Li *et al.* (2007) used descriptions of infections in ancient medical texts to infer the presence of smallpox in China in the 4th century AD; along with other archaeological evidence of smallpox infections, this allowed the authors to place a number of age calibrations at internal nodes in the phylogenetic tree of the variola virus.

Host codivergence – In some cases, the evolutionary timescale of pathogenic organisms can be inferred by assuming correspondence with the evolution of host organisms; for example, the estimated timescale of human movements has been used to provide indirect calibrations for associated pathogens (e.g., Lemey *et al.* 2005). At higher taxonomic levels, the codivergence of endosymbiotic bacteria with their host species has been used for calibrating estimates of long-term rates (Kuo & Ochman 2009). This usually represents the only available source of temporal information for analyses of pathogen evolution across geological scales. Conversely, stably inherited pathogens can

Box 1. *Continued*

be used to infer the recent evolutionary history of their hosts (Kitchen *et al.* 2008). Given that various studies have found evidence of host switching in pathogens (e.g., Ricklefs & Fallon 2002; Harkins *et al.* 2009), the general reliability of assuming host codivergence for age calibration is uncertain.

Heterochronous sampling – In some data sets, such as those containing ancient DNA or serially sampled viruses and bacteria, the sequences have distinct ages. Sometimes these dates are known exactly, for example through museum records or as part of the sampling strategy. Otherwise, they can be estimated radiometrically, stratigraphically or phylogenetically (Ho, Phillips 2009; Shapiro *et al.* 2011). If the age range of the sequences is large in relation to the evolutionary rate, the sampling times can provide sufficient calibrating information for the analysis (Rambaut 2000; Drummond *et al.* 2003). Because they usually apply to the tips of the phylogenetic tree rather than to the internal nodes, calibrations associated with heterochronous sequences differ markedly from those based on palaeontological, geological or archaeological data.

Ancient DNA sequences can be hundreds of thousands of years old (Willerslev *et al.* 2007). In the majority of cases, however, precise dating is bounded at around 50 000 years by the limits of radiocarbon methods. Sample ages can be indirectly obtained by association with dated strata (e.g., Lambert *et al.* 2002), but there is a risk of dating errors because of vertical migration of nucleic acids (Haile *et al.* 2007). Sample intrusions can also produce dating errors, particularly in archaeological contexts.

Heterochronous data can be obtained from viruses or bacteria by serially sampling from a single patient, from different patients over a period of months or years, or from preserved historical samples (e.g., Buonagurio *et al.* 1986; Li *et al.* 1988). The limits on the inclusion of ancient viral and bacterial data are governed by nucleic acid survival, sequence verifiability, access to reliable dating information and the extent of mutational saturation. Virus sequences have been obtained from samples dating to the early twentieth century (e.g., Taubenberger *et al.* 1997), although most data sets comprise samples from the past two or three decades.

Documented pedigrees and laboratory lines – In some studies, such as those of documented pedigrees and laboratory lines, the number of generations separating the individuals is known exactly. If the age of the common ancestor has not been recorded, it can be inferred using an estimate of the mean generation time. Pedigrees can span tens to hundreds of generations, depending on the organism. Meta-analyses of published human pedigree studies can include thousands of genetic transmission events (e.g., Howell *et al.* 2003; Goedbloed *et al.* 2009).

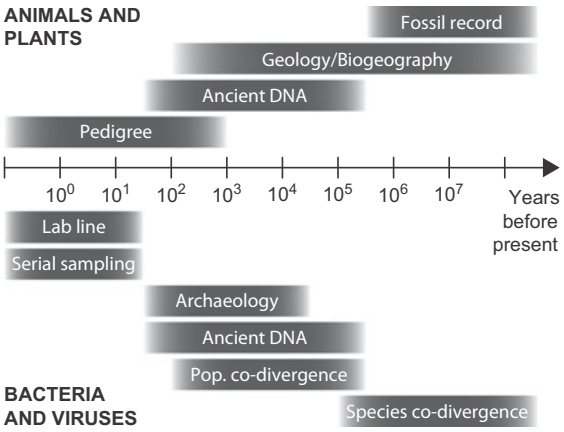


Fig. 2 Typical age ranges of different forms of calibrating information.

Box 2. Evidence for time-dependent molecular rates in animals

Mitochondrial genomes – Perhaps the first observation of time-dependent molecular rates was made in a study of carnivores and primates using gel electrophoresis (Wayne *et al.* 1991), in which the authors noted that the molecular rate showed a linear decline with increasing time depth. This trend was most pronounced for time depths <3 Ma before present. More than a decade later, García-Moreno (2004) collected a number of published estimates of substitution rates for avian mtDNA, finding a distinctive decline in the estimated rate as the calibration point increased in age. Subsequent Bayesian phylogenetic analyses of many of the same avian data sets, along with mitochondrial sequences of primates, reproduced this pattern and clarified the curvilinear relationship (Ho *et al.* 2005, 2007c). Several authors have challenged these findings (Emerson 2007; Bandelt 2008), but a similar rate trend has emerged in a recent study of insect mtDNA (Papadopoulou *et al.* 2010).

The recent work on time-dependent rates followed a series of surprising observations made in pedigree-based studies of humans (Bendall *et al.* 1996; Howell *et al.* 1996, 2003; Mumm *et al.* 1997; Parsons *et al.* 1997; Sigurdardóttir *et al.* 2000; Santos *et al.* 2005). Analyses of mtDNA from individuals with documented relationships yielded remarkably high estimates of mutation rates, with some being more than an order of magnitude higher than estimates of substitution rates based on phylogenetic analyses of mammalian and avian mtDNA. Collectively, these studies pointed towards a very high short-term rate caused by the presence of transient polymorphisms. This raised the question of how long elevated rates might persist, and led to concerns about the accuracy of date estimates for recent evolutionary events (Howell & Mackey 1997; Macaulay *et al.* 1997; Gibbons 1998). High estimates of mtDNA rates have since been obtained in pedigree studies of other organisms, including *Caenorhabditis* (Denver *et al.* 2000), *Drosophila* (Haag-Liautard *et al.* 2008) and Adélie penguins (Millar *et al.* 2008).

In addition to the disparity between pedigree-based and phylogenetic rates, various studies have found that molecular rate estimates are higher when calibrated within species than when calibrated with reference to sister species (e.g., Ho *et al.* 2008; Rajabi-Maham *et al.* 2008; Korsten *et al.* 2009; Davison *et al.* 2011). In some cases, calibration within species was made possible by the inclusion of ancient DNA sequences, which can be up to hundreds of thousands of years old (see Box 1). Most intraspecific ancient mtDNA data sets have produced relatively high estimates of rates, sometimes exceeding long-term phylogenetic rates by an order of magnitude (Lambert *et al.* 2002; Ho *et al.* 2007b, 2011; Kemp *et al.* 2007; Hay *et al.* 2008). In addition, some analyses of ancient mtDNA have found a relationship between estimated rate and calibration age (Ho *et al.* 2007c; Subramanian *et al.* 2009a).

Using calibrations based on dated geological events, such as river capture and lake separation, analyses of mtDNA from cichlids (Genner *et al.* 2007) and galaxiid fishes (Burridge *et al.* 2008) have yielded time-dependent patterns of rates. In both studies, the estimated rate declined over a few hundred thousand years.

There have been several detailed studies of time-dependent rates using human mtDNA. These have been made possible by the availability of complete mtDNA sequences from a large number of modern humans, as well as ancient mtDNA sequences from humans and Neanderthals. Some studies have noted substantial disparities between pedigree-based rates, genealogical rates and phylogenetic rates (Howell *et al.* 2003; Santos *et al.* 2005, 2008; Kemp *et al.* 2007; Endicott & Ho 2008; Endicott *et al.* 2009), while others have provided a more detailed characterization of the dependence of rate on time (Henn *et al.* 2009; Loogväli *et al.* 2009; Soares *et al.* 2009).

Nuclear genomes – Estimates of mutation rates from pedigree studies have indicated a short-term elevation in nuclear DNA which was similar to that seen in mitochondrial DNA. Such estimates have been made for *Caenorhabditis* (Denver *et al.* 2004), *Drosophila* (Houle & Nuzhdin 2004; Haag-Liautard *et al.* 2007; Keightley *et al.* 2009) and human (Nachman & Crowell 2000). These estimates are far higher than substitution rates estimated in phylogenetic analyses.

At the population level, most available nuclear data are in the form of microsatellites or SNPs. Mutation rates vary considerably among microsatellite loci, depending on characteristics such as genomic context, motif size and the existing number of repeats (Ellegren 2000; Buschiazzo & Gemmell 2006; Ballantyne *et al.* 2010). Despite this natural variation, studies of short tandem repeats in the human Y-chromosome have found an approximately threefold

Box 2. *Continued*

difference between the mean rates estimated in pedigree-based and population-level haplogroup-founder analyses (Forster *et al.* 2000; Kayser *et al.* 2000; Zhivotovsky *et al.* 2004, 2006; Vermeulen *et al.* 2009). A smaller disparity was found between a rate estimated from a 13-generation human pedigree and a phylogenetic rate calibrated using the human–chimpanzee divergence (Xue *et al.* 2009).

Box 3. Time-dependent molecular rates in bacteria and viruses

Bacteria – Some degree of time dependence has been observed in rate estimates from bacteria, although there have been few studies of bacterial rates across significant evolutionary timescales. In a study of *Vibrio cholerae*, Feng *et al.* (2008) found that their estimate of the substitution rate, calibrated using heterochronous sequences, was 100-fold higher than the standard rate that is usually assumed. In contrast, slow rates for endosymbiotic bacteria were estimated using calibrations based on the assumption of host codivergence (Kuo & Ochman 2009). The ratio of nonsynonymous to synonymous mutations has been shown to be strongly time-dependent across several closely related bacterial taxa (Rocha *et al.* 2006).

Viruses – Rates of molecular evolution in viruses span several orders of magnitude. This wide variation is a consequence of numerous factors but is primarily influenced by the fidelity of the polymerase that is used during replication (Duffy *et al.* 2008; Holmes 2009). Generally, RNA viruses mutate more rapidly than retroviruses which, in turn, mutate more rapidly than DNA viruses. Nevertheless, there is considerable rate variation within each of these classes, making it difficult to evaluate the dependence of rate on the timescale of observation. This is compounded by the problem of rapid saturation in viral genomes.

Analyses of heterochronous virus samples have typically yielded high estimates of short-term rates, and these differ markedly from rate estimates based on an assumption of host codivergence (Jenkins *et al.* 2002; Ramsden *et al.* 2008; Harkins *et al.* 2009). In a survey of 15 rate estimates from plant viruses, Gibbs *et al.* (2009) found that the highest rate estimates were obtained from heterochronous data. Lower rates were estimated in analyses calibrated at internal nodes, with the very lowest estimates resulting from calibrations based on ancient codivergence events. There also appears to be a discrepancy between intrahost and interhost rates of molecular evolution (e.g., Lemey *et al.* 2006; Gray *et al.* 2011), but this can perhaps be explained by aspects of viral transmission dynamics (Pybus & Rambaut 2009).

There have also been studies of factors leading to the disparity between short- and long-term rates in viruses. There is extensive evidence of strong purifying selection in various viruses (Holmes 2003), with a compelling time-dependent pattern in the ratio of nonsynonymous to synonymous mutations in HIV (Sharp *et al.* 2001). Recent studies have discounted the impact of several potential biases in the high rate estimates obtained from heterochronous viral data (Duffy & Holmes 2009; Firth *et al.* 2010).

rious mutations, which are thought to constitute a large proportion of spontaneous mutations, even in noncoding regions of the genome (Eyre-Walker & Keightley 2007).

To illustrate this point, it is helpful to consider a simple scenario in which we trace the history of a single locus in a pair of sister species in a coalescent framework (Fig. 3). In each species, mutations appear at a certain rate per individual, per generation and per nucleotide. These mutations are eventually lost or fixed (a substitution event) at a rate that depends on their fitness effects and on the population size (Ohta 1973). The fate of neutral mutations in a population is determined solely by genetic drift and, unless there is strong linkage to sites under selection, the fixation rate depends only on the rate at which the mutations are generated (Kimura 1968). At a given population size, the fixation rates of advantageous and deleterious mutations are higher and lower, respectively, than the neutral fixation rate.

In this setting, consider the hierarchy of genetic comparisons that can be made between four contemporary individuals *A*, *B*, *C* and *D* in Fig. 3. *A* and *B* represent closely related individuals from the same population, so the observed genetic differences between them are dominated by mutations, the number of which is determined by the mutation rate. Many of these mutations

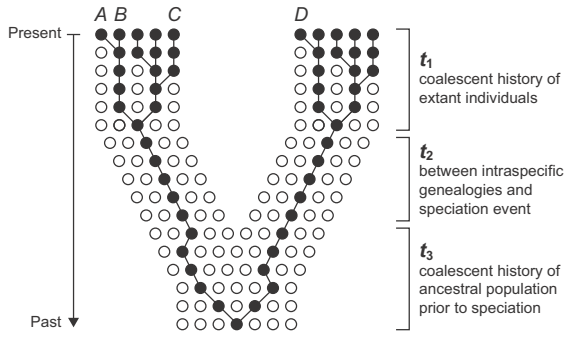


Fig. 3 A simplified representation of genealogical history at a single locus in a pair of species. Each circle represents a randomly mating individual and each row represents one generation. The ancestral species (spanning the time period t_3) splits into two descendent species (spanning the periods t_2 and t_1). Four contemporary individuals labelled *A*, *B*, *C* and *D* are referred to in the text.

are likely to be deleterious, but have not yet been removed by selection. If *A* and *B* are very closely related (e.g., siblings), then a larger proportion of the mutations can be detected and the molecular rate estimate will approach the per replication mutation rate (Howell *et al*. 2003; Santos *et al*. 2005).

A and *C* represent distant relatives within the same population. The number of genetic differences between them is largely determined by the mutation rate but, in the time since their common ancestor, selection has had more opportunity to remove mildly deleterious mutations. For this reason, the estimate of the molecular rate in the *A*–*C* comparison will be considerably lower than that derived from the *A*–*B* comparison.

A and *D* represent individuals from different species. Many of the genetic differences between them are substitutions (mutations that are fixed during t_2; Fig. 3). However, some of the differences might also be polymorphisms that arose during t_1. Additionally, some differences might be due to mutations that occurred before the timing of the speciation event (i.e., during t_3). The relative proportions of each of these periods (t_1, t_2, and t_3) determine the relative influence of the mutation and substitution rates on the molecular rate estimate from the *A*–*D* comparison. If t_2 is long relative to both t_1 and t_3 (e.g., if *A* and *D* are individuals from distantly related species), then the rate estimate will approximate the long-term substitution rate.

The reduction in molecular rate estimates calculated from the *A*–*C* and *A*–*D* comparisons, relative to the *A*–*B* comparison, depends on the effective population size and on the distribution of fitness effects of new mutations in the two lineages. If the population size is large and the majority of mutations are deleterious, then the rate estimated from the *A*–*C* comparison will be considerably lower than from the *A*–*B* comparison. However, if both the population size and the fitness effects of mutations are very small, then the rates estimated from the *A*–*D* and *A*–*C* comparisons may be close to that estimated from the *A*–*B* comparison. Indeed, when all mutations are strictly neutral, the long-term substitution rate is equal to the per generation, per individual mutation rate, regardless of population size (Kimura 1968).

Looking beyond this basic framework, a range of other biological and methodological factors can amplify the disparity between estimates of short- and long-term rates. The identity, importance and relevance of these factors continue to be debated. Below, we broadly class these factors into those that relate to natural selection, calibration error, model misspecification and other artefacts. This classification is by no means definitive; most of the factors involve inadequate modelling of the biological process when estimating molecular evolutionary rates. We describe the factors and discuss their relative significance.

The effects of natural selection

Natural selection tends to remove mutations with negative fitness effects from populations. The ability of negative selection to remove a given mutation depends on both the fitness effect of that mutation and on the effective population size – mutations are removed more efficiently by negative selection when they are highly deleterious and when the effective population size is large (Ohta 1992). Because negative selection removes mutations from populations, it results in long-term rate estimates being lower than short-term rate estimates. The extent to which selection influences the time dependence of rate estimates will depend on the distribution of fitness effects of new mutations. It is possible for positive selection to cause a transient increase in substitution rates, although this will typically apply to only a very small proportion of sites in the genome.

Negative selection

Analyses of molecular data from pedigrees, laboratory mutation-accumulation lines, and intrahost viruses indicate that many new mutations are rapidly removed by selection (Denver *et al.* 2000; Holmes 2003; Haag-Liautard *et al.* 2008; Santos *et al.* 2008; Stewart *et al.* 2008). A large proportion of mutations are likely to be slightly deleterious for the metazoan mitochondrial and viral genomes (Ballard & Whitlock 2004; Duffy *et al.* 2008; Stewart *et al.* 2008; Galtier *et al.* 2009), the primary sources of evidence for time-dependent rates (see Boxes 2 and 3). If this is the case, we should expect to see a rapid decline from spontaneous mutation rates as

we look back in time. However, Hill–Robertson interference, whereby a low rate of recombination prevents multiple selected sites from evolving independently (Hill & Robertson 1966), can substantially reduce the efficacy of selection in some sequences (Williamson & Orive 2002), consequently tempering the effects of negative selection on the time dependence of molecular rate estimates.

The effect of negative selection on molecular rate estimates is also apparent in the time-dependent decline in the ratio of nonsynonymous to synonymous changes in protein-coding sequences (Sharp *et al.* 2001; Elson *et al.* 2004; Rocha *et al.* 2006; Peterson & Masel 2009; Subramanian 2009). Analyses of mitochondrial DNA (mtDNA) sequences have shown that rates at nonsynonymous sites display stronger time dependence than those at synonymous sites (Endicott & Ho 2008; Loogväli *et al.* 2009; Soares *et al.* 2009; Subramanian *et al.* 2009a). These results suggest an important role for negative selection in the time dependence of rates. However, theoretical investigations have indicated that the action of purifying selection alone is insufficient to explain the time-dependent patterns of rates obtained in published empirical studies of primates and birds (Woodhams 2006; Peterson & Masel 2009). Sequences in noncoding regions, such as the mitochondrial D-loop and microsatellites in the human Y-chromosome, have also been shown to exhibit time-dependent rates (Box 2). This pattern might be the result of negative selection on these sequences, but it is perhaps more likely that it is because of negative selection in closely linked coding regions.

Positive selection

Positive selection can produce an increase in nonsynonymous rates, which is reflected in the overall substitution rate as well as the ratio of nonsynonymous to synonymous rates. Positive selection could explain some observations of time dependence if there has been recent adaptive evolution. It has been postulated that some of the nonsynonymous mutations towards the tips of the human mitochondrial genealogy represent adaptive changes in response to climatic variation (Mishmar *et al.* 2003; Ruiz-Pesini *et al.* 2004), although these claims have been contested (Elson *et al.* 2004; Sun *et al.* 2007). Moreover, it has been suggested that this form of positive selection does not lead to the persistent elevation of rates seen in empirical studies (Loogväli *et al.* 2009). Given the wide range of sequences and taxa in which time-dependent rates have been detected, it seems implausible that widespread recent increases in rates of adaptive molecular evolution could offer a common explanation.

The effects of calibration errors

To estimate absolute rates of molecular evolution, it is necessary to use independent temporal information to calibrate the age of a given node in the tree (Box 1). If calibrations are specified incorrectly, any resulting rate estimates might be biased. A number of errors can arise when choosing and implementing calibrations; most of these errors lead to overestimates of rates. This overestimation bias tends to be largest on short timescales, which can lead to a time-dependent pattern of rates (as seen in Fig. 1).

Coalescent calibration error

The genetic divergence of a given locus always precedes or coincides with population or species divergences (t_3 in Fig. 3), unless there is subsequent gene flow. When calibrations are based on the presumed timing of a population- or species-divergence event, genetic divergence is typically assumed to coincide with population divergence. This leads to an underestimate of the time since genetic divergence, with a consequent overestimation of substitution rates. The magnitude of this bias will be greatest on short timescales, where the temporal difference between genetic and population divergences forms an appreciable proportion of the total time separating the two populations or species (Edwards & Beerli 2000). The severity of the effect also increases with the effective population size and generation time of the ancestral lineage. Some methods enable calibration of population divergences rather than genetic divergences (e.g., Heled & Drummond 2010), allowing the problem of coalescent calibration error to be avoided.

Recently, it was suggested that the rate overestimation caused by this coalescent calibration error provides a sufficient explanation for observed time-dependent trends in rates (Peterson & Masel 2009). In practice, however, there are some instances in which congruence between genetic and population splits is plausible. If there had been a bottleneck in the ancestral population prior to or during speciation, then there is a high probability that genetic divergence occurred close to the time of population divergence.

Coalescent calibration error almost certainly explains some portion of the observed time dependence of molecular rate estimates. However, the suggestion that it offers a sufficient explanation is challenged by the high rate estimates obtained in studies of data sets comprising sequences of varying ages (heterochronous data), including ancient mtDNA (e.g., Lambert *et al.* 2002; Ho *et al.* 2007b; Hay *et al.* 2008) and viruses (Box 3). In these studies, rate estimates are not cali-

brated at divergence events, but at the tips of the tree (see Box 1). Therefore, this approach does not entail any assumptions about the concurrence of genetic and population divergences. If the high rates estimated in analyses of heterochronous data are legitimate, then coalescent calibration error cannot be the exclusive cause of time-dependent rates. Therefore, it appears that heterochronous sequence data might play an important role in further explication of this phenomenon, although there has been some debate over the validity of rates estimated from such data sets (Ho *et al.* 2007b, 2011; Debruyne & Poinar 2009; Miller *et al.* 2009; Navascués & Emerson 2009; Subramanian *et al.* 2009b; Firth *et al.* 2010).

Palaeontological, biogeographic and archaeological calibration error

In phylogenetic analyses, estimates of molecular rates are usually calibrated using temporal information from the fossil record or from assumed biogeographic splits (Box 1). The timing of an evolutionary divergence event can be assumed to be older than the earliest appearance of either of its descendent lineages in the fossil record. Genetic divergence generally precedes the appearance of diagnostic morphological variation, meaning that there is a disparity between date estimates provided by molecular data and the fossil record. In addition, there is also a very low probability that a fossil taxon can be reliably placed close to a divergence event. Thus, genetic divergence times are underestimated by palaeontological evidence, which can lead to overestimates of molecular rates. The relative magnitude of this bias is most problematic for studies of short evolutionary timescales where the age underestimation can represent a substantial proportion of the true age of the calibration (Ho *et al.* 2005), but fossil calibrations are rarely used in such settings.

In studies of recent evolutionary events, calibrations based on biogeography or archaeology are sometimes employed (Box 1); for example, the appearance of a geological feature can produce a barrier to gene flow, allowing the timing of genetic divergence to be estimated. In some cases, colonization events can be identified in phylogenetic trees and their timing estimated using geological or archaeological evidence (e.g., Fleischer *et al.* 1998; Henn *et al.* 2009). Biogeographic calibrations carry a number of risks because they usually involve strong assumptions, the most important of which is that there is a close correspondence between genetic and geological events. However, genetic divergence can substantially antedate the emergence of a barrier to gene flow, for example in the presence of undetected, extinct sister lineages (Emerson 2007). This can lead to a severe overestimation of molecular rates (e.g., Marko 2002). On the other hand, dispersal or other agents of gene flow can result in genetic divergence postdating geological events, resulting in an overestimation of the calibration age and subsequent underestimation of molecular rates.

The effect of model misspecification

In all estimates of rates, numerous simplifying assumptions are required. This is reflected in the use of relatively simple models of genetic inheritance, nucleotide substitution and demographic history. Computational and statistical tractability is the main purpose of these assumptions, but often this comes at the cost of biological realism. Biases in the estimates of various parameters, including rates, can arise if the model is misspecified (Swofford *et al.* 2001; Lemmon & Moriarty 2004). Although this problem can sometimes be minimized through a rigorous model-selection procedure, it can be the case that none of the available models is entirely adequate (Gatesy 2007).

Phylogenetic assumptions concerning inheritance

In most phylogenetic analyses, it is assumed that the sampled sequences are orthologous, which means that they diverged from each other via a speciation process and not by duplication events or horizontal gene transfer. In analyses of animal mtDNA, common assumptions include uniparental inheritance, homoplasmy and a lack of recombination. However, occurrences of biparental inheritance, paternal leakage, heteroplasmy and recombination have been reported for mtDNA from a variety of taxa (e.g., Schwartz & Vissing 2002; Bromham *et al.* 2003; Barr *et al.* 2005; Santos *et al.* 2005; Tsaousis *et al.* 2005). If these violations of the assumptions are not adequately dealt with, estimates of the rate of molecular evolution can be inaccurate (White *et al.* 2008). The presence of nontarget sequences (e.g., paralogous copies of genes in a single-gene analysis) will usually produce an artificial inflation of genetic diversity, leading to a rate overestimation that is greater on short timescales. However, it is difficult to envisage these problems occurring on a widespread scale in published analyses.

Substitution model misspecification and saturation

When the process of nucleotide substitution is modelled in a phylogenetic analysis, it is almost always assumed to be homogeneous throughout the tree. One of the functions of these models is to correct for unobserved substitutions; underestimation of these unseen changes

is the typical outcome of substitution model misspecification (Sullivan & Joyce 2005). Under-correction for mutational saturation can cause patterns of time-dependent rates, because saturation is likely to be less of a problem over very short time frames (such as those associated with analyses of recent events) but is an important factor over longer time frames (such as those associated with phylogenetic analyses). It has been suggested that this effect might be at least partly responsible for time-dependent rates (García-Moreno 2004; Emerson 2007; Ho *et al.* 2007c; Raquin *et al.* 2008).

Demographic factors

In Bayesian phylogenetic methods, a prior distribution needs to be specified for the tree (including topology and branch lengths). In some implementations, the prior distribution is generated using a stochastic branching or coalescent process (Drummond *et al.* 2006; Yang & Rannala 2006). For intraspecific data sets, the use of the coalescent also requires assumptions about the demographic history of the species being studied. Coalescent-based approaches typically involve simple parametric models of demographic change, built upon assumptions such as panmixia and random sampling. It has been demonstrated that violation of these assumptions, as well as incorrect modelling of demographic history, can lead to biases in rate estimates; for example, the presence of severe bottlenecks or marked population subdivision, combined with biased sampling, can artificially elevate the estimates of short-term rates made using heterochronous data (Miller *et al.* 2009; Navascués & Emerson 2009).

The problem of demographic model misspecification can be at least partly addressed using extensions of the coalescent that allow for population structure (Notohara 1990), while a model-selection procedure can be implemented to determine the most appropriate description of population history. Alternatively, some complexities in demographic history can be accommodated using flexible models of population change (e.g., Drummond *et al.* 2005; Heled & Drummond 2008). However, because population history can rarely be described with confidence, it remains possible that time-dependent rates will be a feature of coalescent analyses for some time.

Artefacts causing time dependence of molecular rates

Time dependence in rate estimates can also be influenced by a number of additional measurement errors. Some of these, such as sequence errors, can be difficult to detect, whereas certain statistical phenomena can be dealt with more readily.

Sequence error and postmortem damage

Errors in DNA sequences can be introduced at various stages of analysis, including base misincorporation during the polymerase chain reaction and misinterpretation of sequence traces. In studies of sequences acquired from ancient samples, the DNA molecule can be altered by postmortem damage, the level of which is usually higher than typical sequencing error by an order of magnitude (Lindahl 1993). Regardless of the source, sequence errors will manifest themselves as spurious recent mutations, leading to time-dependent rate estimates as well as biases in other parameters (Clark & Whittam 1992; Ho *et al.* 2005, 2007a; Johnson & Slatkin 2008). The manifold sequencing coverage achieved by next-generation sequencers might lead to a better characterization of sequence quality, but the degree of error in published sequences is unknown (e.g., Forster 2003; Bandelt *et al.* 2006).

The impact of sequence errors is most pronounced when there is low diversity among sequences, which is often the case for population-level data. Assuming a fixed frequency of sequence error, the estimation bias in rates declines with increasing sequence divergence (Ho *et al.* 2005). In phylogenetic analyses of ancient DNA sequences, this estimation bias can be alleviated using models of sequence damage (Ho *et al.* 2007a; Mateiu & Rannala 2008; Rambaut *et al.* 2009). However, because all sequencing errors will be incorrectly recorded as recent mutations, they are likely to explain some proportion of the time dependence of rates in a large number of studies.

Skewed rate distributions

Molecular rates are described by scale parameters, which have a natural lower bound of zero and lack a rigid upper bound. This asymmetric constraint can introduce a bias in Bayesian rate estimates, particularly if they are associated with wide credibility intervals, leading to a positively skewed posterior distribution in which the point probability of a very high rate is higher than normal. Consequently, the mean can give a biased reflection of the rate, especially if the rate is estimated from a data set with very low sequence variability (Heled & Drummond 2008; Debruyne & Poinar 2009). Thus, the use of mean rate estimates in Bayesian studies of substitution rates might explain some cases of time-dependent rates. A recent simulation-based study demonstrated that this bias can be reduced or removed through the use of other measures of central tendency, such as the median or mode (Ho *et al.* 2011).

Ascertainment bias

For some types of markers, comparison of rates estimated over different timescales might be confounded if the rates have been calculated using different loci. This can be caused by ascertainment bias, whereby rapidly mutating markers are selected for pedigree studies while slowly evolving markers are chosen for evolutionary studies (Zhivotovsky *et al.* 2004; Bandelt 2008; Ballantyne *et al.* 2010). This can also be the case for single-nucleotide polymorphisms (SNPs), whereby highly variable sites are selected to provide resolution within species but might not provide resolution in interspecific comparisons; for example, this is a major problem with SNP-chips that have been designed to analyse a small subset of species (Decker *et al.* 2009). In some studies, fast-evolving and poorly aligned sites are discarded to improve the ratio of phylogenetic signal to noise and to increase confidence in homology (Castresana 2000), leading to a biased representation of the sequence data. Although it is unclear whether these ascertainment biases extend to any of the direct comparisons that have been made between pedigree and phylogenetic rates (Howell *et al.* 2008), the focus on rapidly evolving markers does lead to the additional problem of mutational saturation over long time frames.

Concluding remarks

Numerous factors have the potential to produce a time-dependent pattern in molecular rates, particularly via elevation of short-term rates. Short-term estimates of rates are likely to include mutations that will not be observed on longer timescales, owing to loss of mutations by selection. Therefore, short-term rates are closer to the nonlethal mutation rate than long-term rates, with the latter reflecting slower substitution rates.

In addition to the transience of deleterious mutations, various estimation biases can lead to an inflation of short-term measures of rates. In general, estimation biases and methodological artefacts tend to cause overestimation of short-term rates and underestimation of long-term rates. Overestimation bias tends to be proportionally greater on short timescales, which leads to a time-dependent pattern of rates. Although it is likely that their contributions will vary considerably among data sets, particularly among taxonomic groups, it seems plausible that negative selection, coalescent calibration error and perhaps model misspecification are the most important factors leading to elevated short-term rates. Some of these issues can be corrected by rigorous testing of model assumptions and careful selection of calibrating information.

Given the number of potential factors causing time-dependent rates, as well as differences in their effects among loci, it might prove very difficult to correct for these factors when making estimates of evolutionary timescales. Until further studies have quantitatively addressed the importance of and interactions between each of these factors, building mechanistic models to account for time-dependent rates is unlikely to be straightforward. Perhaps the most pragmatic approach to obtaining accurate estimates of evolutionary timescales is to be aware of the factors causing time-dependent rates and to attempt to avoid them.

An important practical consequence of time-dependent rates is that a strict molecular clock is rarely appropriate for investigating phylogenetic trees that cross the barrier between populations and species. Even if a relaxed molecular clock is employed in an attempt to deal with this problem (e.g., Korsten *et al.* 2009), it is typically assumed that all of the branch-specific rates come from a single distribution (Drummond *et al.* 2006). A potential solution is to perform separate analyses at the intra- and interspecific levels, or perhaps to use local-clock models to distinguish between short- and long-term rates (Yoder & Yang 2000; Drummond & Suchard 2010). Alternatively, it may be appropriate to define substitution rate mathematically as a function of time, applied across all lineages of a phylogenetic tree (Rodrigo *et al.* 2008).

Further empirical studies, including detailed comparisons among the factors discussed in this review, will help to identify the scale and impact of the various factors. Evidence from nuclear data, including noncoding regions, should also prove to be illuminating. The growth of genomic sequence data provides an excellent opportunity to investigate time-dependent rates in detail, particularly when sequences are available from multiple individuals within a population as well as from closely related species. The identification of reliable age calibrations, particularly those relating to recent evolutionary timescales, is also an important avenue of research.

It should be noted that some of the factors described in this review can also affect estimates of rates over much longer time frames; for example, the problems of saturation and model misspecification will grow in severity as the timescale of measurement increases. On long timescales, however, it is likely that the effects of time-dependent factors will be small in comparison with other sources of error, including those relating to calibrations and models of rate variation among lineages. Additional factors such as changes in the substitution process and differences in equilibrium nucleotide frequencies will also become significant (e.g., Phillips 2009).

Understanding the causes of time-dependent rates is important for a number of reasons, not only for quanti-

fying rates of molecular change but also for constructing evolutionary timescales. This is likely to have an impact on estimates of the timing of events that have occurred in the past few hundred thousand years, and perhaps over longer timescales. Disentangling the factors influencing rate estimation will also lead to an improvement in our understanding of the molecular evolutionary process.

Acknowledgements

We thank the Genetics Society of Australasia for sponsoring the 2008 symposium in Adelaide that stimulated this review. We thank Jack da Silva for helpful discussions and Chris Burridge and two anonymous reviewers for constructive feedback. SYWH, RL, LB, MJP, and AC were funded by the Australian Research Council. SYWH was also supported by a start-up grant from the University of Sydney. JS was supported by the University of Adelaide.

References

Ballantyne KN, Goedbloed M, Fang R *et al.* (2010) Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. *American Journal of Human Genetics*, **87**, 341–353.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

Bandelt HJ (2008) Time dependency of molecular rate estimates: tempest in a teacup. *Heredity*, **100**, 1–2.

Bandelt H-J, Kong QP, Richards M, Macaulay V (2006) Estimation of mutation rates and coalescence times: some caveats. In:*Human Mitochondrial DNA and the Evolution of* Homo sapiens (eds Bandelt H-J, Macaulay V, Richards M), pp. 47–90. Springer, Berlin.

Barr CM, Neiman M, Taylor DR (2005) Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. *New Phytologist*, **168**, 39–50.

Bendall KE, Macaulay VA, Baker JR, Sykes BC (1996) Heteroplasmic point mutations in the human mtDNA control region. *American Journal of Human Genetics*, **59**, 1276–1287.

Bromham L, Eyre-Walker A, Smith NH, Maynard Smith J (2003) Mitochondrial Steve: paternal inheritance of mitochondria in humans. *Trends in Ecology and Evolution*, **18**, 2–4.

Buonagurio DA, Nakada S, Parvin JD *et al.* (1986) Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in NS gene. *Science*, **232**, 980–982.

Burridge CP, Craw D, Fletcher D, Waters JM (2008) Geological dates and molecular rates: fish DNA sheds light on time dependency. *Molecular Biology and Evolution*, **25**, 624–633.

Buschiazzo E, Gemmell NJ (2006) The rise, fall and renaissance of microsatellites in eukaryotic genomes. *Bioessays*, **28**, 1040–1050.

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, **17**, 540–552.

Clark AG, Whittam TS (1992) Sequencing errors and molecular evolutionary analysis. *Molecular Biology & Evolution*, **9**, 744–752.

Davison J, Ho SYW, Bray SC *et al.* (2011) Late-quaternary biogeography scenarios for the brown bear (*Ursus arctos*), a wild mammal model species. *Quaternary Science Reviews*, **30**, 418–430.

Debruyne R, Poinar HN (2009) Time dependency of molecular rates in ancient DNA data sets, a sampling artifact? *Systematic Biology*, **58**, 348–360.

Decker JE, Pires JC, Conant GC *et al.* (2009) Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. *Proceedings of the National Academy of Sciences, USA*, **106**, 18644–18649.

Denver DR, Morris K, Lynch M, Vassilieva LL, Thomas WK (2000) High direct estimate of the mutation rate in the mitochondrial genome of *Caenorhabditis elegans*. *Science*, **289**, 2342–2344.

Denver DR, Morris K, Lynch M, Thomas WK (2004) High mutation rate and predominance of insertions in the *Caenorhabditis elegans* nuclear genome. *Nature*, **430**, 679–682.

Drummond AJ, Suchard MA (2010) Bayesian random local clocks, or one rate to rule them all. *BMC Biology*, **8**, 114.

Drummond AJ, Pybus OG, Rambaut A, Forsberg R, Rodrigo AG (2003) Measurably evolving populations. *Trends in Ecology and Evolution*, **18**, 481–488.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22**, 1185–1192.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, e88.

Duffy S, Holmes EC (2009) Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *Journal of General Virology*, **90**, 1539–1547.

Duffy S, Shackelton LA, Holmes EC (2008) Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews Genetics*, **9**, 267–276.

Edwards SV, Beerli P (2000) Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839–1854.

Ellegren H (2000) Heterogeneous mutation processes in human microsatellite DNA sequences. *Nature Genetics*, **24**, 400–402.

Elson JL, Turnbull DM, Howell N (2004) Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. *American Journal of Human Genetics*, **74**, 229–238.

Emerson BC (2007) Alarm bells for the molecular clock? No support for Ho et al.'s model of time-dependent molecular rate estimates. *Systematic Biology*, **56**, 337–345.

Endicott P, Ho SYW (2008) A Bayesian evaluation of human mitochondrial substitution rates. *American Journal of Human Genetics*, **82**, 895–902.

Endicott P, Ho SYW, Metspalu M, Stringer C (2009) Evaluating the mitochondrial timescale of human evolution. *Trends in Ecology and Evolution*, **24**, 515–521.

Eyre-Walker A, Keightley PD (2007) The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, **8**, 610–618.

Feng L, Reeves PR, Lan R *et al.* (2008) A recalibrated molecular clock and independent origins for the cholera pandemic clones. *PLoS ONE*, **3**, e4053.

Firth C, Kitchen A, Shapiro B *et al.* (2010) Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Molecular Biology and Evolution*, **27**, 2038–2051.

Fleischer RC, McIntosh CE, Tarr CL (1998) Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K-Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. *Molecular Ecology*, **7**, 533–545.

Forster P (2003) To err is human. *Annals of Human Genetics*, **67**, 2–4.

Forster P, Röhl A, Lünnemann P *et al.* (2000) A short tandem repeat–based phylogeny for the human Y chromosome. *American Journal of Human Genetics*, **67**, 182–196.

Galtier N, Nabholz B, Glemin S, Hurst GD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology*, **18**, 4541–4550.

García-Moreno J (2004) Is there a universal mtDNA clock for birds? *Journal of Avian Biology*, **35**, 465–468.

Gatesy J (2007) A tenth crucial question regarding model use in phylogenetics. *Trends in Ecology and Evolution*, **22**, 509–510.

Genner MJ, Seehausen O, Lunt DH *et al.* (2007) Age of cichlids: new dates for ancient lake fish radiations. *Molecular Biology and Evolution*, **24**, 1269–1282.

Gibbons A (1998) Calibrating the mitochondrial clock. *Science*, **279**, 28–29.

Gibbs AJ, Fargette D, García-Arenal F, Gibbs MJ (2009) Time – the emerging dimension of plant virus studies. *Journal of General Virology*, **91**, 13–22.

Gingerich PD (1983) Rates of evolution: effects of time and temporal scaling. *Science*, **222**, 159–161.

Gingerich PD (2001) Rates of evolution on the time scale of the evolutionary process. *Genetica*, **112–113**, 127–144.

Goedbloed M, Vermeulen M, Fang RN *et al.* (2009) Comprehensive mutation analysis of 17 Y-chromosomal short tandem repeat polymorphisms included in the AmpFlSTR Yfiler PCR amplification kit. *International Journal of Legal Medicine*, **123**, 471–482.

Gray RR, Parker J, Lemey P *et al.* (2011) The mode and tempo of hepatitis C virus evolution within and among hosts. *BMC Evolutionary Biology*, **11**, 131.

Haag-Liautard C, Dorris M, Maside X *et al.* (2007) Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature*, **445**, 82–85.

Haag-Liautard C, Coffey N, Houle D *et al.* (2008) Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*. *PLoS Biology*, **6**, e204.

Haile J, Holdaway R, Oliver K *et al.* (2007) Ancient DNA chronology within sediment deposits: are paleobiological reconstructions possible and is DNA leaching a factor? *Molecular Biology and Evolution*, **24**, 982–989.

Harkins GW, Delport W, Duffy S *et al.* (2009) Experimental evidence indicating that mastreviruses probably did not co-diverge with their hosts. *Virology Journal*, **6**, 104.

Hay JM, Subramanian S, Millar CD, Mohandesan E, Lambert DM (2008) Rapid molecular evolution in a living fossil. *Trends in Genetics*, **24**, 106–109.

Heads M (2005) Dating nodes on molecular phylogenies: a critique of molecular biogeography. *Cladistics*, **21**, 62–78.

Heled J, Drummond AJ (2008) Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, **8**, 289.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–580.

Henn BM, Gignoux CR, Feldman MW, Mountain JL (2009) Characterizing the time dependency of human mitochondrial DNA mutation rate estimates. *Molecular Biology and Evolution*, **26**, 217–230.

Hill WG, Robertson A (1966) The effect of linkage on limits to artificial selection. *Genetical Research*, **8**, 269–294.

Ho SYW, Larson G (2006) Molecular clocks: when times are a-changin'. *Trends in Genetics*, **22**, 79–83.

Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**, 367–380.

Ho SYW, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution*, **22**, 1561–1568.

Ho SYW, Heupink TH, Rambaut A, Shapiro B (2007a) Bayesian estimation of sequence damage in ancient DNA. *Molecular Biology and Evolution*, **24**, 1416–1422.

Ho SYW, Kolokotronis S-O, Allaby RG (2007b) Elevated substitution rates estimated from ancient DNA. *Biology Letters*, **3**, 702–705.

Ho SYW, Shapiro B, Phillips M, Cooper A, Drummond AJ (2007c) Evidence for time dependency of molecular rate estimates. *Systematic Biology*, **56**, 515–522.

Ho SYW, Saarma U, Barnett R, Haile J, Shapiro B (2008) The effect of inappropriate calibration: three case studies in molecular ecology. *PLoS ONE*, **3**, e1615.

Ho SYW, Lanfear R, Phillips MJ *et al.* (2011) Bayesian estimation of substitution rates from ancient DNA sequences with low information content. *Systematic Biology*, **60**, 366–375.

Holmes EC (2003) Patterns of intra- and interhost nonsynonymous variation reveal strong purifying selection in dengue virus. *Journal of Virology*, **77**, 11296–11298.

Holmes EC (2009) *The Evolution and Emergence of RNA Viruses*. Oxford University Press, Oxford, UK.

Houle D, Nuzhdin SV (2004) Mutation accumulation and the effect of *copia* insertions in *Drosophila melanogaster*. *Genetics Research*, **83**, 7–18.

Howell N, Mackey D (1997) Reply to Macauley et al. *American Journal of Human Genetics*, **61**, 986–990.

Howell N, Kubacka I, Mackey DA (1996) How rapidly does the human mitochondrial genome evolve? *American Journal of Human Genetics*, **59**, 501–509.

Howell N, Smejkal CB, Mackey DA *et al.* (2003) The pedigree rate of sequence divergence in the human mitochondrial genome: there is a difference between phylogenetic and pedigree rates. *American Journal of Human Genetics*, **72**, 659–670.

Howell N, Howell C, Elson JL (2008) Time dependency of molecular rate estimates for mtDNA: this is not the time for wishful thinking. *Heredity*, **101**, 107–108.

Jenkins GM, Rambaut A, Pybus OG, Holmes EC (2002) Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *Journal of Molecular Evolution*, **54**, 156–165.

Johnson PL, Slatkin M (2008) Accounting for bias from sequencing error in population genetic estimates. *Molecular Biology and Evolution*, **25**, 199–206.

Kayser M, Roewer L, Hedman M *et al.* (2000) Characteristics and frequency of germline mutations at microsatellite loci

from the human Y chromosome, as revealed by direct observation in father/son pairs. *American Journal of Human Genetics*, **66**, 1580–1588.

Keightley PD, Trivedi U, Thomson M *et al.* (2009) Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Research*, **19**, 1195–1201.

Kemp BM, Malhi RS, McDonough J *et al.* (2007) Genetic analysis of early holocene skeletal remains from Alaska and its implications for the settlement of the Americas. *American Journal of Physical Anthropology*, **132**, 605–621.

Kimura M (1968) Evolutionary rate at the molecular level. *Nature*, **217**, 624–626.

Kitchen A, Miyamoto MM, Mulligan CJ (2008) Utility of DNA viruses for studying human host history: case study of JC virus. *Molecular Phylogenetics and Evolution*, **46**, 673–682.

Korsten M, Ho SYW, Davison J *et al.* (2009) Sudden expansion of a single brown bear maternal lineage across northern continental Eurasia after the last ice age: a general demographic model for mammals? *Molecular Ecology*, **18**, 1963–1979.

Kuo C-H, Ochman H (2009) Inferring clocks when lacking rocks: the variable rates of molecular evolution in bacteria. *Biology Direct*, **4**, 35.

Kurtén B (1959) Rates of evolution in fossil mammals. *Cold Spring Harbor Symposia on Quantitative Biology*, **24**, 205–215.

Lambert DM, Ritchie PA, Millar CD *et al.* (2002) Rates of evolution in ancient DNA from Adélie penguins. *Science*, **295**, 2270–2273.

Lemey P, Pybus OG, Van Dooren S, Vandamme AM (2005) A Bayesian statistical analysis of human T-cell lymphotropic virus evolutionary rates. *Infection, Genetics and Evolution*, **5**, 291–298.

Lemey P, Rambaut A, Pybus OG (2006) HIV evolutionary dynamics within and among hosts. *AIDS Reviews*, **8**, 125–140.

Lemmon AR, Moriarty EC (2004) The importance of proper model assumption in Bayesian phylogenetics. *Systematic Biology*, **53**, 265–277.

Li W, Tanimura M, Sharp P (1988) Rates and dates of divergence between AIDS virus nucleotide sequences. *Molecular Biology & Evolution*, **5**, 313–330.

Li Y, Carroll DS, Gardner SN *et al.* (2007) On the origin of smallpox: correlating variola phylogenics with historical smallpox records. *Proceedings of the National Academy of Sciences, USA*, **104**, 15787–15792.

Lindahl T (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.

Loogväli E-L, Kivisild T, Margus T, Villems R (2009) Explaining the imperfection of the molecular clock of hominid mitochondria. *PLoS ONE*, **4**, e8260.

Macaulay VA, Richards MB, Forster P *et al.* (1997) mtDNA mutation rates – no need to panic. *American Journal of Human Genetics*, **61**, 983–985.

Marko PB (2002) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Molecular Biology and Evolution*, **19**, 2005–2021.

Mateiu LM, Rannala BH (2008) Bayesian inference of errors in ancient DNA caused by postmortem degradation. *Molecular Biology and Evolution*, **25**, 1503–1511.

Millar CD, Dodd A, Anderson J *et al.* (2008) Mutation and evolutionary rates in Adélie penguins from the Antarctic. *PLoS Genetics*, **4**, e1000209.

Miller HC, Moore JA, Allendorf FW, Daugherty CH (2009) The evolutionary rate of tuatara revisited. *Trends in Genetics*, **25**, 13–15.

Mishmar D, Ruiz-Pesini E, Golik P *et al.* (2003) Natural selection shaped regional mtDNA variation in humans. *Proceedings of the National Academy of Sciences, USA*, **100**, 171–176.

Mumm S, Whyte MP, Thakker RV, Buetow KH, Schlessinger D (1997) mtDNA analysis shows common ancestry in two kindreds with X-linked recessive hypoparathyroidism and reveals a heteroplasmic silent mutation. *American Journal of Human Genetics*, **60**, 153–159.

Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in humans. *Genetics*, **156**, 297–304.

Navascués M, Emerson BC (2009) Elevated substitution rate estimates from ancient DNA: model violation and bias of Bayesian methods. *Molecular Ecology*, **18**, 4390–4397.

Notohara M (1990) The coalescent and the genealogical process in geographically structured population. *Journal of Mathematical Biology*, **29**, 59–75.

O'Fallon BD, Seger J, Adler FR (2010) A continuous-state coalescent and the impact of weak selection on the structure of gene genealogies. *Molecular Biology and Evolution*, **27**, 1162–1172.

Ohta T (1973) Slightly deleterious mutant substitutions in evolution. *Nature*, **246**, 96–98.

Ohta T (1992) The nearly neutral theory of molecular evolution. *Annual Review of Ecology and Systematics*, **23**, 263–286.

Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Molecular Biology and Evolution*, **27**, 1659–1672.

Parsons TJ, Muniec DS, Sullivan K *et al.* (1997) A high observed substitution rate in the human mitochondrial DNA control region. *Nature Genetics*, **15**, 363–368.

Penny D (2005) Relativity for molecular clocks. *Nature*, **426**, 183–184.

Peterson GI, Masel J (2009) Quantitative prediction of molecular clock and K_a/K_s at short timescales. *Molecular Biology and Evolution*, **26**, 2595–2603.

Phillips MJ (2009) Branch-length estimation bias misleads molecular dating for a vertebrate mitochondrial phylogeny. *Gene*, **441**, 132–140.

Pybus OG, Rambaut A (2009) Evolutionary analysis of the dynamics of viral infectious disease. *Nature Reviews Genetics*, **10**, 540–550.

Rajabi-Maham H, Orth A, Bonhomme F (2008) Phylogeography and postglacial expansion of *Mus musculus domesticus* inferred from mitochondrial DNA coalescent, from Iran to Europe. *Molecular Ecology*, **17**, 627–641.

Rambaut A (2000) Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics*, **16**, 395–399.

Rambaut A, Ho SYW, Drummond AJ, Shapiro B (2009) Accommodating the effect of ancient DNA damage on inferences of demographic histories. *Molecular Biology and Evolution*, **26**, 245–248.

Ramsden C, Melo FL, Figueiredo LM *et al.* (2008) High rates of molecular evolution in hantaviruses. *Molecular Biology and Evolution*, **25**, 1488–1492.

Raquin AL, Depaulis F, Lambert A *et al.* (2008) Experimental estimation of mutation rates in a wheat population with a gene genealogy approach. *Genetics*, **179**, 2195–2211.

Ricklefs RE, Fallon SM (2002) Diversification and host switching in avian malaria parasites. *Proceedings of the Royal Society B*, **269**, 885–892.

Rocha EP, Smith JM, Hurst LD *et al.* (2006) Comparisons of dN/dS are time dependent for closely related bacterial genomes. *Journal of Theoretical Biology*, **239**, 226–235.

Rodrigo A, Bertels F, Heled J *et al.* (2008) The perils of plenty: what are we going to do with all these genes? *Philosophical Transactions of the Royal Society London Series B*, **363**, 3893–3902.

Roopnarine PD (2003) Analysis of rates of morphologic evolution. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 605–632.

Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC (2004) Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science*, **303**, 223–226.

Santos C, Montiel R, Sierra B *et al.* (2005) Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: a model using families from the Azores Islands (Portugal). *Molecular Biology & Evolution*, **22**, 1490–1505.

Santos C, Montiel R, Arruda A *et al.* (2008) Mutation patterns of mtDNA: empirical inferences for the coding region. *BMC Evolutionary Biology*, **8**, 167.

Schwartz M, Vissing J (2002) Paternal inheritance of mitochondrial DNA. *New England Journal of Medicine*, **347**, 576–580.

Shapiro B, Ho SYW, Drummond AJ *et al.* (2011) A Bayesian method to estimate unknown sequence ages in a phylogenetic context. *Molecular Biology and Evolution*, **28**, 879–887.

Sharp PM, Bailes E, Chaudhuri RR *et al.* (2001) The origins of acquired immune deficiency syndrome viruses: where and when? *Philosophical Transactions of the Royal Society London Series B*, **356**, 867–876.

Sigurdardóttir S, Helgason A, Gulcher JR, Stefánsson K, Donnelly P (2000) The mutation rate in the human mtDNA control region. *American Journal of Human Genetics*, **66**, 1599–1609.

Soares P, Ermini L, Thomson N *et al.* (2009) Correcting for purifying selection: an improved human mitochondrial molecular clock. *American Journal of Human Genetics*, **84**, 1–20.

Stewart JB, Freyer C, Elson JL *et al.* (2008) Strong purifying selection in transmission of mammalian mitochondrial DNA. *PLoS Biology*, **6**, e10.

Subramanian S (2009) Temporal trails of natural selection in human mitogenomes. *Molecular Biology and Evolution*, **26**, 715–717.

Subramanian S, Denver DR, Millar CD *et al.* (2009a) High mitogenomic evolutionary rates and time dependency. *Trends in Genetics*, **25**, 482–486.

Subramanian S, Hay JM, Mohandesan E, Millar CD, Lambert DM (2009b) Molecular and morphological evolution in tuatara are decoupled. *Trends in Genetics*, **25**, 16–18.

Sullivan J, Joyce P (2005) Model selection in phylogenetics. *Annual Review of Ecology and Systematics*, **36**, 445–466.

Sun C, Kong QP, Zhang YP (2007) The role of climate in human mitochondrial DNA evolution: a reappraisal. *Genomics*, **89**, 338–342.

Swofford DL, Waddell PJ, Huelsenbeck JP *et al.* (2001) Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Systematic Biology*, **50**, 525–539.

Taubenberger JK, Reid AH, Krafft AE, Bijwaard KE, Fanning TG (1997) Initial genetic characterization of the 1918 ''Spanish'' influenza virus. *Science*, **275**, 1793–1796.

Tsaousis AD, Martin DP, Ladoukakis ED, Posada D, Zouros E (2005) Widespread recombination in published animal mtDNA sequences. *Molecular Biology and Evolution*, **22**, 925–933.

Vermeulen M, Wollstein A, van der Gaag K *et al.* (2009) Improving global and regional resolution of male lineage differentiation by simple single-copy Y-chromosomal short tandem repeat polymorphisms. *Forensic Science International: Genetics*, **3**, 205–213.

Wayne RK, van Valkenburgh B, O'Brien SJ (1991) Molecular distance and divergence time in carnivores and primates. *Molecular Biology & Evolution*, **8**, 297–319.

White DJ, Wolff JN, Pierson M, Gemmell NJ (2008) Revealing the hidden complexities of mtDNA inheritance. *Molecular Ecology*, **17**, 4925–4942.

Willerslev E, Cappellini E, Boomsma W *et al.* (2007) Ancient biomolecules from deep ice cores reveal a forested southern Greenland. *Science*, **317**, 111–114.

Williamson S, Orive ME (2002) The genealogy of a sequence subject to purifying selection at multiple sites. *Molecular Biology and Evolution*, **19**, 1376–1384.

Woodhams M (2006) Can deleterious mutations explain the time dependency of molecular rate estimates? *Molecular Biology and Evolution*, **23**, 2271–2273.

Xue Y, Wang Q, Long Q *et al.* (2009) Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. *Current Biology*, **19**, 1453–1457.

Yang Z, Rannala B (2006) Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution*, **23**, 212–226.

Yoder AD, Yang ZH (2000) Estimation of primate speciation dates using local molecular clocks. *Molecular Biology & Evolution*, **17**, 1081–1090.

Zhivotovsky LA, Underhill PA, Cinnioglu C *et al.* (2004) The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *American Journal of Human Genetics*, **74**, 50–61.

Zhivotovsky LA, Underhill PA, Feldman MW (2006) Difference between evolutionarily effective and germ-line mutation rate due to stochastically varying haplogroup size. *Molecular Biology and Evolution*, **23**, 2268–2270.

S.Y.W.H. is interested in molecular clocks, evolutionary timescales, and ancient DNA. R.L. and L.B. study the causes and consequences of molecular evolution. M.J.P. uses phylogenetic inference to understand evolutionary and ecological processes. J.S. is a doctoral candidate investigating how evolutionary timescales could be refined in the light of ancient DNA and improved molecular clock methods. A.R. is the director of NESCent and develops computational evolutionary genetic methods to model the processes that shape the diversity of rapidly evolving pathogens. A.C. is the director of the Australian Centre for Ancient DNA at the University of Adelaide.

Chapter 5

Study of the time dependence of molecular rates

(Part 2)

The influence of rate heterogeneity among sites on the time dependence of molecular rates

Statement of authorship

The influence of rate heterogeneity among sites on the time dependence of molecular rates

Julien Soubrier (Candidate)

Contributed to the design of the research program; performed analyses, interpreted the data and created figures; co-wrote the manuscript and acted as corresponding author.

I hereby certify that the statement of contribution is accurate.

Signed .... *Date* ...02.04.12...

Mike Steel

Contributed to research concept; designed and wrote the mathematical part; helped in data interpretation, figure development and manuscript evaluation.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .. .. *Date 29 November 2011*

Mike Lee

Contributed to research concept and program; helped with data interpretation and manuscript development and evaluation.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .... *Date* ...6.2.12...

Clio Der Sarkissian

Assisted with design of mathematical analysis, data interpretation, figure development and manuscript evaluation.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ..... *Date* ...28/11/11...

Stéphane Guindon

Developed the 'FreeRate' option of PhyTime, helped with designing research for the PhyTime part and provided comments on, and evaluation of the manuscript.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed…….. …….. *Date* 5/12/2012

Simon Y. W. Ho

Contributed to research concept and program; supervised development of work, helped in data interpretation and manuscript development and writing.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… *Date* ……05.04.2012……

Alan Cooper

Conceived the research program; supervised development of work, data interpretation and manuscript development and writing.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed . …… *Date* ……

Soubrier, J., Steel, M., Lee, M.S.Y., Sarkissian, C.D., Guindon, S., Ho, S.Y.W. & Cooper, A. (2012)
The influence of rate heterogeneity among sites on the time dependence of molecular rates.
Molecular Biology and Evolution, v. 29(11), pp. 3345-3358

Chapter 6

Ancient whole mitochondrial genomes to study recent human evolution

(Part 1)

The mystery of human mitochondrial haplogroup H and the genetic origins of Europeans

Statement of authorship

The mystery of human mitochondrial haplogroup H and the genetic origins of Europeans

Paul Brotherton
Devised the library preparation and target enrichment protocol, co-developed the research concept, co-wrote the manuscript and acted as corresponding author.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …. …… *Date* 8/4/12

Wolfgang Haak
Developed the research concept, processed the samples, analysed data, co-wrote the manuscript, and acted as corresponding author.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ….. …… *Date* 02/04/2012

Jennifer Templeton
Established and optimised probe preparation and target enrichment protocols.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .. …… *Date* 02/04/2012

Julien Soubrier (Candidate)
Performed the Bayesian skyride analysis and wrote the corresponding supplementary information paragraph. Participated to the figure development.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed …… …… *Date* 02.04.12

Christina J Adler
Analysed data, contributed corresponding methods chapters.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .. …… *Date* 17/4/2012

Stephen Richards
Performed experimental steps for SMRT sequencing and provided supplementary information paragraph.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .. *Date* 02/04/2012

Guido Brandt
Provided ancient samples, replication of results, and contextual information.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed *Date* 13.04.12

Robert Ganslmeier
Provided archaeological and contextual information.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ... . *Date* .....05.04.2012.....

Mannis van Oven
Provided phylogenetic analyses and access to unpublished data.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .. .. *Date* 01 Mar 2012

Rosalie Kenyon
Prepared Mitochip v2.0 runs and analysed data.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .... *Date* .....02/03/12.....

Mark Van der Hoek
Prepared Mitochip v2.0 runs and analysed data.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .... *Date* ...2-3-12.....

Jonas Korlach
Provided access to Pacific Biosciences platform, analysed SMRT sequencing data and contributed corresponding methods chapter.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .... *Date* ..3/1/12.....

Khai Luong
Analysed SMRT sequencing data and provided corresponding methods chapter.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .... *Date* ...3/1/12.....

Simon Y. W. Ho
Supervised the Bayesian skyride setup and interpretation.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .... *Date* .....05.04.2012.....

Lluis Quintana-Murci
Provided critical unpublished data from modern-day populations.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ... . *Date 12 Apr. 12*

Doron M Behar
Provided critical unpublished data from modern-day populations.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed *Date 18 Apr 2012*

Harald Meller
Co-developed the underlying concept (diachronic approach) and provided archaeological and contextual information.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ..... *Date* 13.03.2012

Kurt W Alt
Co-developed the underlying concept (diachronic approach) and provided ancient samples and contextual information.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ..... *Date March 15, 2012*

Alan Cooper
Co-developed the concept and experimental design, and co-wrote the manuscript.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .. *Date* 05.04.2012

The mystery of human mitochondrial haplogroup H and the genetic origins of Europeans

Paul Brotherton[1, §,*], Wolfgang Haak[1*], Jennifer Templeton[1], Guido Brandt[2], Julien Soubrier[1], Christina J Adler[1], Stephen M Richards[1], Robert Ganslmeier[3], Mannis van Oven[4], Rosalie Kenyon[5], Mark B Van der Hoek[5], Jonas Korlach[6], Khai Luong[6], Simon Y W Ho[7], Lluis Quintana-Murci[8], Doron M Behar[9], Harald Meller[3], Kurt W Alt[2], Alan Cooper[1], & The Genographic Consortium[10]

[1]The Australian Centre for Ancient DNA, University of Adelaide, Adelaide, South Australia 5005, Australia.

[2]Institute of Anthropology, Colonel-Kleinmann Weg 2, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany.

[3]State Office for Heritage Management and Archaeology Saxony-Anhalt / State Museum for Prehistory Halle, Richard-Wagner-Straße 9, D-06114 Halle/Saale, Germany.

[4]Department of Forensic Molecular Biology, Erasmus MC, University Medical Center Rotterdam, 3000 CA Rotterdam, The Netherlands.

[5]SA Pathology, Adelaide, South Australia 5000, Australia.

[6]Pacific Biosciences, USA.

[7]School of Biological Sciences, University of Sydney, New South Wales 2006, Australia

[8]Institut Pasteur, Paris, France.

[9]Rambam Medical Center, 31096 Haifa, Israel.

[10]Consortium members are listed as Supplementary Information

[§] New address: Archaeogenetics Research Group, School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, UK.

[*]These authors contributed equally to this manuscript

Corresponding authors:

Dr Paul Brotherton
Archaeogenetics Research Group
School of Applied Sciences
University of Huddersfield
Queensgate
Huddersfield HD1 3DH
UK.
Ph: +44 1484-471676
E-mail: P.M.Brotherton@hud.ac.uk

Dr Wolfgang Haak
Australian Centre for Ancient DNA
School of Earth & Environmental Sciences
North Terrace Campus
The University of Adelaide
SA-5005, Adelaide
Australia
Ph: +61 8 8303 5565
Fax: +61 8 8303 4364
E-mail: wolfgang.haak@adelaide.edu.au

Dr Alan Cooper
Australian Centre for Ancient DNA
School of Earth & Environmental Sciences
North Terrace Campus
The University of Adelaide
SA-5005, Adelaide
Australia
Ph: +61 8 8303 5950
Fax: +61 8 8303 4364
E-mail: alan.cooper@adelaide.edu.au

ABSTRACT

The archaeological record of Central Europe identifies a succession of profound cultural and economic changes between the last hunter-gatherers of the Mesolithic and the incoming farmers of the Early Neolithic (~5450 BC), through to the socially stratified chiefdoms of the Early Bronze Age (~2200 BC)[1-3]. The exact nature and genetic context of the transformative changes that took place over these four millennia remains unclear[4,5], although current genetic patterns hint at multiple inputs from outside Central Europe[6,7]. Present-day European mitochondrial (mt) DNA variability is dominated by haplogroup (hg) H (~40%)[6], yet this was virtually absent in Mesolithic hunter-gatherers[8] and present in only 19% of early Neolithic farmers[9]. Traditional mitochondrial D-loop sequencing has been unable to accurately resolve the evolutionary history of hg H, leaving the origins of more than a third of European mtDNA diversity obscured[10]. We use new ancient DNA approaches to sequence 39 complete hg H mt genomes from well-preserved human remains and reconstruct the genetic background to the cultural changes that took place between the Neolithic and Bronze Age in 'real-time'. This first population-level study of ancient human mt genomes allows us to directly observe sequence evolution in human mtDNA by linking individual hg H genomes to specific points in time. We calculate an evolutionary rate that is considerably higher than current estimates and reconstruct recent demographic history. The ancient mt genomes reveal that the origins of present-day European hg H distribution and diversity were established by the Mid-Neolithic, shortly after a pronounced genetic discontinuity between archaeological cultures of the Early (pre-4000 BC) and Mid-to-Late Neolithic (post-4000 BC). Further genetic contributions were made in the Late Neolithic, particularly by the first pan-European cultures such as the Bell Beakers who expanded out of Iberia ca. 2800 BC, potentially in association with the initial spread of the Celtic language family[11].

Mitochondrial haplogroup H accounts for over 40% of mtDNA variation across much of Western Eurasia, with declining frequencies south and east to ~10–30% in the Near East and Caucasus[6]. Phylogeographic studies suggest that hg H arrived in Europe from the Near East prior to the Last Glacial Maximum (22,000 BP), and survived in glacial refugia in Southwest Europe before undergoing a postglacial re-expansion[7,12]. However, traditional approaches, which have relied on sequencing ~360 bp of the mitochondrial D-loop or control region, have been unable to resolve either the phylogeny or phylogeographic distribution of hg H[6,13]. Limited studies of complete genome sequences have highlighted the complexity of the evolutionary history of this haplogroup (e.g.[12,14-16]). As a result, the origins of almost half of the maternal population history of Europe remain unclear, raising the question of *when* and *how* H became the predominant haplogroup.

A recent phylogenetic analysis of the 16.6 kb human mt genome has identified 87 subclades of hg H[17] and revealed that 71% of hg H polymorphic diversity is located outside the D-loop, in the coding region[18]. Consequently, we designed a new DNA-hybridisation capture system to sequence complete mt genomes from immortalised libraries of the highly damaged and degraded endogenous DNA recovered from archaeological remains (Supplementary Information)[19].

We used this method to generate a genetic picture across a chronological transect of archaeological cultures in the Mittelelbe-Saale region of Saxony-Anhalt (Germany). This region provides a continuous archaeological record since Palaeolithic times with excellent preservation of human skeletal remains. Our transect spans the >3500 years of the Central European Neolithic period (Table 1; Supplementary Information), from the first farmers of the Early Neolithic *linear pottery culture* LBK (5450-4775 BC), through the Rössen (4625-4250 BC), Schöningen (4100-3950 BC), Baalberge (3950-3400 BC), and Salzmünde (3400-3025 BC) cultures. These were followed by two of the first pan-European phenomena, the Corded Ware (CWC, 2800-2050 BC) and Bell Beaker (BBC, 2500-2050 BC) complexes, before the emergence of the Early Bronze Age with the Unetice culture (2200-1575 BC).

Individuals assigned to hg H by simplex and multiplex PCR[9] were selected from a collection of over 400 prehistoric Europeans (Supplementary Information). Target-enriched DNA libraries were prepared from 37 Mittelelbe-Saale individuals, as well as two hg H samples from Italy, and genotyped via the Affymetrix MitoChip v2.0[20] (Supplementary Information). Six target-enriched libraries were also analysed via a single-molecule, real-time (SMRT[21]; Pacific Biosciences) sequencing platform. In addition, 35/391 (9%) of all SNPs identified via the MitoChip were independently tested and confirmed by direct PCR from independent extracts and Sanger sequencing.

Table 1. Summary of genotyping and archaeological data for human individuals.
All samples displayed sequence variants at nucleotide positions (np) 263, 750, 1438, 4769, 8860 and 15326 - going 'out of hg H2a2a1/rCRS' - except for BZH8 (missing 1438 and 4769 due to its position within hg H2a1a; Figure 1a). Bold = previously characterised hg H SNPs. Regular= 'private' / as-yet-unknown hg H SNPs (nps 16519 is usually not considered in phylogenetic reconstructions).

Culture / Age	Hg[1]	SNP differences relative to the rCRS	Individual[2]
LBK (5450-4775 BC)	**H23**	**10211**, 16519	HAL36
	H	16093, 16129, 16519	HAL11
	H26	**11152**, 16519	HAL32
	H1e	**3010**, **5460**, 16519	HAL39
	H88	**8596**, 16519	**DEB9**
	H1j	**3010**, **4733**, 16519	DEB21
	H1bz	1719, **3010**, 14380, 16519	KAR6a
	H	152, 16519	KAR11b
	H46b	**2772**, 11893, 16519	KAR16a
Rössen (4625-4475/4250 BC)	**H89**	6932, 8068, 12696, 16519	**OSH2**
	H1	**3010**, 16519	OSH3
	H16	**152**, **10394**, 16519	OSH1
	H5b	**456**, **5471**, **16304**	OSH7
Schöningen (4100-3950 BC)	**H10i**	13503, **14470A**, **16093**, 16519	SALZ18a
	H1e7	1766, **3010**, **5460**, 16519	SALZ21b
Baalberge (3950-3400 BC)	**H1e1a5**	**3010**, **5460**, (5960), **8512**, 8865, **14902**, 16519	ESP30
	H7d5	**4793**, **15409**, 16388, 16519	HQU4
Salzmünde (3400-3100/3025 BC)	**H3**	152, **6776**, 16519	SALZ57a
	H3	**6776**, 16519	SALZ77a
Corded Ware (2800-2200//2050 BC)	**H6a1a**	**239**, **3915**, **4727**, **9380**, **11253**, **16362**, **16482**	ESP15
	H1 TBD	**3010**, 8149, 9377, 9467, 13671, 14319, **16189**, 16519	**BZH6**
Bell Beaker (2500-2200/2050 BC)	**H1e7**	**3010**, **5460**, 15220, 15401, 16293, 16519	**BZH4**
	H5a3	**456**, **513**, **4336**, **15884**, **16304**	ROT6
	H3b	**2581**, **6776**, 16519	ALB1
	H3ao2	4577, **6776**, **16256**, 16519	ROT1
	H5a3	**456**, **513**, **4336**, **15884**, **16304**	ROT2
	H4a1	**3992**, **4024**, **5004**, **9123**, **14365**, **14582**	QUEXII1
	H4a1	**3992**, **4024**, **5004**, **9123**, **14365**, **14582**	QUEXII2
	H1	**3010**, 16519	QLB26a
	H13a1a2c	**2259**, **4745**, 9025, **13542**, **13680**, **14872**	**QUEXII3**
	H1	**3010**, 16519	QLB28b
Unetice (2200-1575 BC)	**H11a**	**195**, **961G**, **8448**, **(13759)**, **16293**, **16311**	**BZH1**
	H2a1a3	**951**, **6173**, **13095**, 16240T, **16354**	BZH8
	H82a	195, **16220**, 16519	BZH14
	H4a1a1a5	**73**, **3992**, **4024**, **5004**, **8269**, **9123**, **10044**, 13545, **14365**, **14582**	EUL41a
	H3	152, **6776**, 16519	EUL57B
	H7h	**4793**, **16213**,16519	QUEVIII4
Nuragic Bronze Age (1624 BC)	**H1aw1**	**3010**, **8701**, 15912, 16519	-
Iron Age (500 BC)	**H90**	5435, 8911, 10237, 15109, 16519	-

[1]Haplogroup H designations based on the http://www.phylotree.org mtDNA tree Build 14 (5 April 2012)[17,18].
[2]Samples that were independently typed via Pacific Bioscience SMRT sequencing are shown in bold.

Mt genome sequences from all 39 individuals were unambiguously assignable to sub-branches of hg H[18], confirming that a single human individual was typed in each case (Table 1). The ancient hg H sequences were highly diverse, with 34 distinct haplotypes attributed to 20 major sub-haplogroups (gene diversity H= 0.997 +/- 0.0071; nucleotide diversity 0.000421 +/- 0.000225) including three novel lineages (H88-H90).

Phylogenetic network analysis of these ancient mt genomes reveals clear evidence of evolution over the 3500-year transect (Figure 1a). Sequences from older samples represent basal hg H lineages, only one to three mutations away from the ancestral root haplotype of hg H, while more recent samples comprise more derived haplotypes, and appear on longer branches. This represents the first detailed direct observation of the human mt evolutionary rate and also confirms the authenticity of the aDNA data. The network also reveals pronounced differences between the sub-haplogroup composition of mt genomes from the Early Neolithic (ENE) cultures and those of the Mid/Late Neolithic to Early Bronze Age (LNE). Early Neolithic (and in particular LBK) genomes from before 4000 BC, are either rare today (H16, H23, H26) or have not yet been observed in present-day populations (H46b, H88, H89). In sharp contrast, later hg H sub-clades (after 4000 BC) are largely common in present-day European populations (e.g., hg H3, H4, H6, H7, H11, and H13)[12,14,15]. Only three of 37 lineages (within the common, basal sub-clade H1) were shared between ENE and LNE cultures, demonstrating minimal local genetic continuity (Table 1).

To examine this apparent temporal discontinuity, we tested whether cultures from early and late Neolithic pools represent different meta-populations. While population pairwise distances were non-significant (F_{ST} = 0.01459; p = 0.12048+/- 0.0032), comparisons based on the presence of sub-clades using non-parametric multivariate analysis of variance (NP-MANOVA, p=0.0034) confirmed the pattern (Supplementary Tables 8 and 9). Therefore, it appears that Early Neolithic lineages were largely superseded after 4000 BC, and that much of present-day Central Europe's hg H diversity was created by the diversification of Mid-Neolithic lineages (after 4000 BC) and/or by the incorporation of new lineages from outside central Europe during the subsequent Late Neolithic (2800 BC) and emerging Bronze Age (2200 BC). The Late Neolithic is known to have been a period of profound cultural and economic change[3] with newly emerging, pan-European cultures, such as the Bell Beaker phenomenon in Western Europe and the Corded Ware culture in north eastern Europe, overlapping in Mittelelbe-Saale (Supplementary Information). These demographic changes between Early and Late Neolithic cultures parallel the initial Meso-Neolithic transition, the other major genetic boundary – between hunter-gatherers and farmers in early European history – thus far revealed by ancient DNA[8,9].

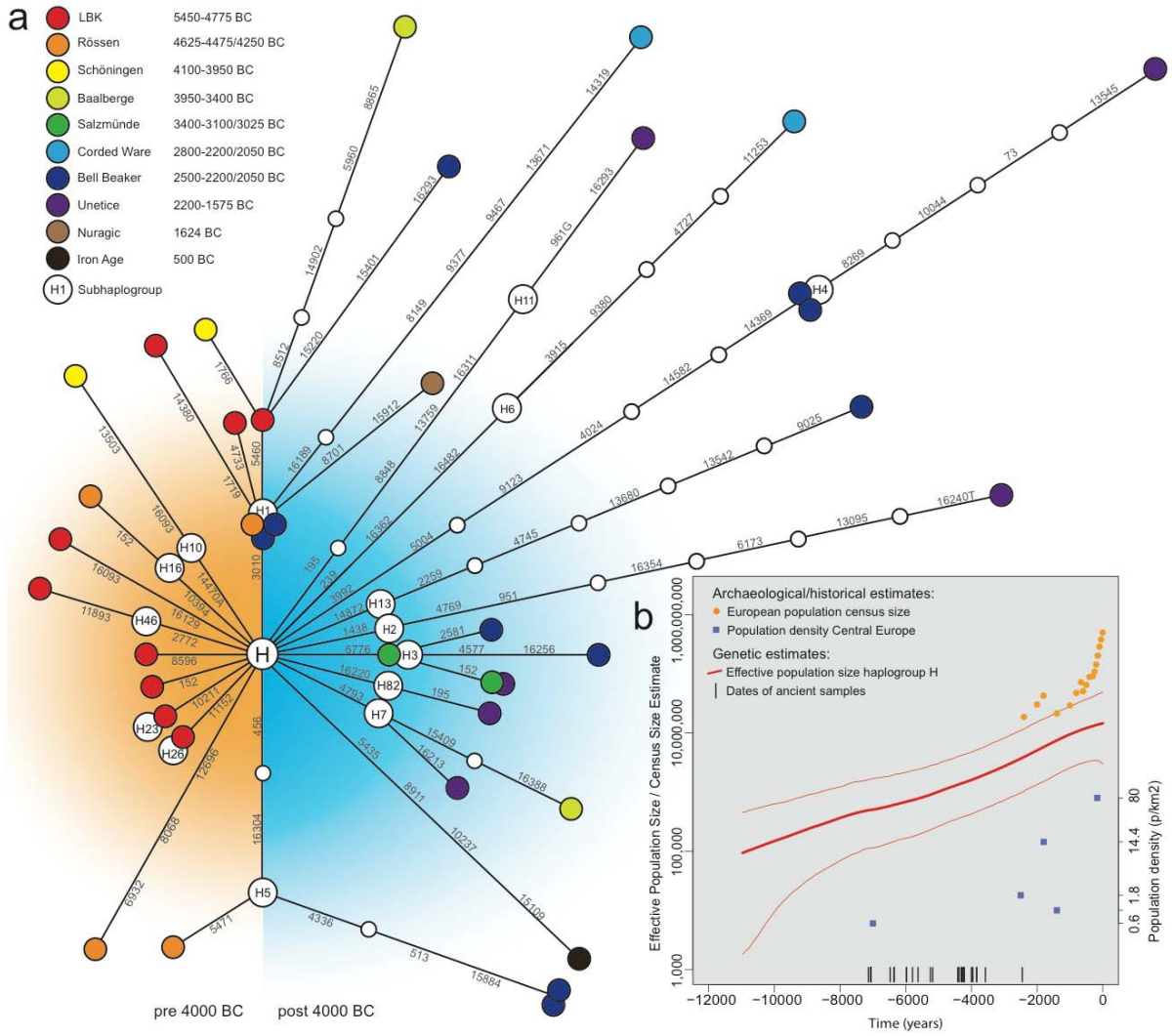


Figure 1. Mitochondrial haplogroup H sequence evolution. **a**, A phylogenetic network of 39 prehistoric mitochondrial genomes illustrates rapid genetic change during the Neolithic. The sorting of mt genomes into two temporal groupings (pre- and post-4000 BC) reveals a disjunct sub-haplogroup distribution between Early Neolithic and Mid-to-Late Neolithic, respectively. Node colours represent archaeological cultures. **b**, A Bayesian skyride plot of 200 representative present-day and 39 ancient hg H mt genomes demonstrates consistent exponential growth of the effective population size during the entire Holocene (thick red line denotes the posterior median, thinner flanking lines denote the 95% credibility interval; note the logarithmic scale of the y-axes). Prehistoric samples (18 radiocarbon and 21 mean archaeological dates) served as internal calibration points (black bars). For comparison, census size estimates for the European population are shown as orange dots. Population density estimates from the archaeological record for key periods in Central Europe are plotted as blue squares in chronological order: LBK, Iron Age, Roman period, Merovingian, and Pre-industrial modern times (y-axis on the right).

It is unclear when hg H became the predominant haplogroup in Europe, as archaeogenetic and paleodemographic reconstructions using modern data have very large uncertainties[7,22]. We used Bayesian skyride analysis to reconstruct the population history of hg H using 200 representative present-day samples and the 39 dated ancient specimens as calibration points (Figure 1b; Supplementary Information). The resulting skyride plot is the first real-time paleodemographic reconstruction of European populations and shows a consistent and strong exponential growth in effective population size through the entire Holocene. While the data match European census size[23] and population density estimates from archaeological sites[24], they also provide detailed estimates for prehistoric times for which data points remain very scarce (Figure 1b). As predicted by the temporal dependency of evolutionary rates[25], the rate obtained from the dated ancient DNA sequences appears considerably higher (2.4×10^{-8} substitutions/site/year for the entire mt genome) than standard estimates based on the traditional human/chimp split (*e.g.* 1.66×10^{-8} for the entire mt genome[26] and 1.26×10^{-8} for the coding region[27]). Consequently, we estimate a younger coalescence date for hg H (10.9 to 19.1 kya) than previously reported (19.2 to 21.4 kya for HVSI[6], 15.7 to 22.5 kya for the mt coding region[27] or 14.7-22.6 kya when corrected for purifying selection[26]). This suggests that the re-expansion of hg H lineages into Central, Northern, and Eastern Europe reflects population growth throughout the entire Neolithic period, rather than the traditional view of just a post-glacial expansion prior to the Holocene (12 kya)[6,7,12].

To examine the geographic origin of Neolithic cultures (Table 1) and their contribution to the modern Central European mtDNA pool, we used Principal Component Analysis (PCA) to examine their genetic affinities to present-day Western Eurasian populations (Supplementary Table 7). A PCA of 15 hg H subgroup frequencies shows that present-day populations form three significantly supported ($p<0.0001$) geographic clusters: (i) Iberia in the west; (ii) the Caucasus, the Near East, and Anatolia; and (iii) (Central and Eastern) Europe from the Urals to France (Figure 2; Supplementary Tables 7, 8a). Discriminant analysis of the first 15 principal components scores confirms that the clustering pattern of the modern and prehistoric cultures is non-random ($p<0.001$; Supplementary Figure 6).

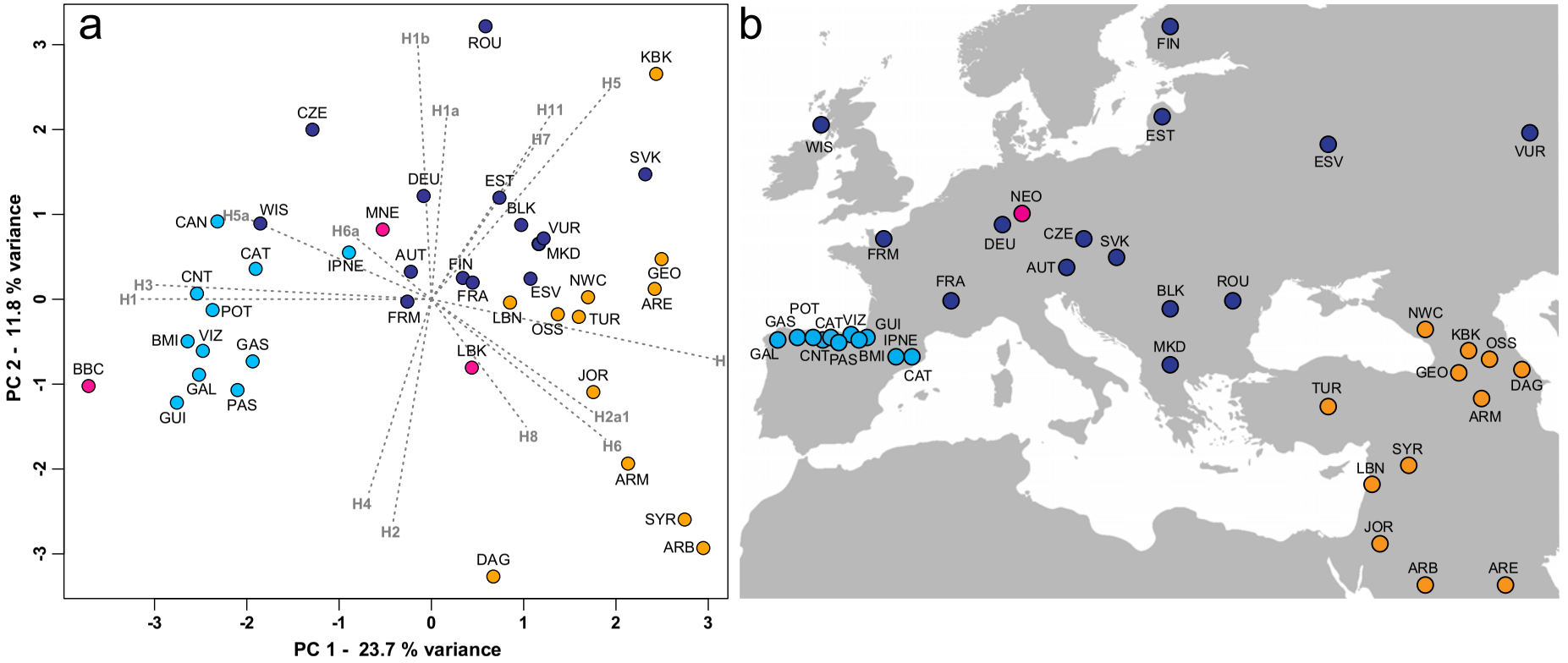


Figure 2. Population affinities of select Neolithic cultures. **a,** PCA plot based on frequencies of 15 hg H sub-haplogroups from 37 present-day Western Eurasian populations. **b,** Map of Europe depicting the geographic origins of the present-day samples (light blue: Western Europe; dark blue: Central and Eastern Europe; orange; Near East, Caucasus, and Anatolia) and ancient samples (pink). Populations are abbreviated as follows (Supplementary Table 7): GAL, Galicia; CNT, Cantabria; CAT, Catalonia; GAS, Galicia/Asturia; CAN, Cantabria2; POT, Potes; PAS, Pasiegos; VIZ, Vizcaya; GUI, Guipuzcoa; BMI, Basques, IPNE, Iberian Peninsula Northeast; TUR, Turkey; ARM, Armenia; GEO, Georgia; NWC, Northwest Caucasus; DAG, Dagestan; OSS, Ossetia; SYR, Syria; LBN, Lebanon; JOR, Jordan; ARB, Arabian Peninsula; ARE, Arabian Peninsula2; KBK, Karachay-Balkaria; MKD, Macedonia; VUR, Volga-Ural region; FIN, Finland; EST, Estonia; ESV, Eastern Slavs; SVK, Slovakia; FRA, France; BLK, Balkans, DEU, Germany; AUT, Austria, ROU, Romania; FRM, France Normandy; WIS, Western Isles; CZE, Czech Republic; LBK, Linear pottery culture; BBC, Bell Beaker culture; MNE, Middle Neolithic. Discriminant analysis of the first 15 principal components scores revealed that the clustering pattern of the modern and prehistoric cultures was non-random ($p<0.001$; Supplementary Figure 6). The three distinct clusters of present-day populations were significantly ($p<0.0001$) supported by a NP-MANOVA of the 15 hg H subgroup frequencies (Supplementary Table 8a).

The PCA plot reveals that earliest farmers (LBK; n=9) cluster with present-day Caucasus, Near East, and Anatolian populations, as previously noted[9] and consistent with a detailed archaeological record tracing the spread of agriculture via LBK farmers back to Anatolia and the Near East where farming originated ~12,000 years ago[1].

In contrast, individuals from the successor series of regional (Mid-)Neolithic cultures (Rössen, Schöningen, Baalberge, and Salzmünde, ca. 4625-3025 BC, MNE; n=10) cluster with present-day Central European populations (Figure 2). This is consistent with a phase of more localised development after the initial Neolithisation (Supplementary Information), even allowing for moderate influences from contemporaneous Mid-Neolithic cultures outside Mittelelbe-Saale. Importantly, this indicates substantial genetic continuity from the Mid-Neolithic to the present-day in Central Europe.

The Late Neolithic Bell Beakers (BBC; n=7) form one of the first pan-European archaeological complexes, emerging around 2800 BC, and are recognised by a cultural package of rich grave goods (including the eponymous bell-shaped ceramic beakers). BBC individuals display close genetic affinities to present-day Iberian populations (Figure 2), largely based on high frequencies of sub-clades H1 and H3, which are thought to have spread from a post-glacial Iberian refugium[7,10,12] and were also found in ancient Neolithic samples from France and Spain[28,29]. These genetic affinities between Central Europe's Bell Beakers and present-day Iberian populations suggest a considerable genetic influx from the West during the Late Neolithic. The role of the Bell Beaker folk in the rise and spread of hg H across Western Europe also has intriguing linguistic implications, as this movement has recently been linked to the initial spread of the Celtic language family[11]. This theory suggests that the Celtic language family developed from Indo-European precursors in Iberia (e.g. *Tartessian*) and spread throughout the Atlantic Zone, before a period of rapid mobility reflected by the Beaker phenomenon carried them across much of Western Europe. It now seems possible that this movement also led to the high proportion of hg H1 and H3 in the same areas. This important finding not only challenges traditional views of a linguistic spread of Celtic westwards from Central Europe during the Iron Age, but also implies that Indo-European languages arrived in Europe substantially earlier, presumably with the arrival of farming[30].

Additional Late Neolithic movements added further genetic complexity, as exemplified by the eastern affinities of the Corded Ware culture (CWC). The archaeological association of the CWC is supported by two distinct mt genomes (H1_TBD and H6a1a), which were not found in previous cultures and the contemporaneous Bell Beaker neighbours. Similarly, data from the subsequent Early Bronze Age Unetice culture indicate genetic

influences from both the East (sub-clades H2a, H7 and H11) and the West (sub-clades H3 and H4) based on frequency distributions of these sub-clades in present-day populations[12,14,15].

Our results suggest that the broad foundations of the Central European mtDNA pool, approximated via hg H, were formed during the Neolithic and that through a series of later Neolithic demographic events, hg H became the predominant mtDNA branch in Europe. Early Neolithic hg H mt lineages brought in by Central Europe's first farmers were largely superseded after 4000 BC, but since that time there appears to have been substantial genetic continuity from the Mid-Neolithic to the present-day in Central Europe. The widespread Bell Beaker cultural phenomenon, potentially connected to the expansion of the Celtic language family, is likely to have been one driving factor in the later expansion of hg H. However, Late Neolithic Corded Ware and early Bronze Age data suggest a complex series of further genetic contributions.

References

1 Whittle, A. W. R. & Cummings, V. *Going over: the mesolithic-neolithic transition in North-West Europe*. (Oxford University Press, 2007).

2 Sherratt, A. in *Patterns of the Past: Studies in honour of David Clarke* (eds I. Hodder, G. Isaac, & N. Hammond) 261-305 (Cambridge University Press, 1981).

3 Heyd, V. Families, prestige goods, warriors & complex societies: Beaker groups of the 3rd millennium cal BC along the upper & middle Danube. *P Prehist Soc* **73**, 327-379 (2007).

4 Ammerman, A. J. & Cavalli-Sforza, L. L. *The neolithic transition and the genetics of populations in Europe*. (Princeton University Press, 1984).

5 Whittle, A. *Europe in the Neolithic. The Creation of New Worlds*. (Cambridge University Press, 1996).

6 Richards, M. *et al.* Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* **67**, 1251-1276, doi:S0002-9297(07)62954-1 [pii] (2000).

7 Soares, P. *et al.* The archaeogenetics of Europe. *Curr Biol* **20**, R174-183, doi:S0960-9822(09)02069-710.1016/j.cub.2009.11.054 (2010).

8 Bramanti, B. *et al.* Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science (New York, N.Y.)* **326**, 137-140, doi:10.1126/science.1176869 (2009).

9 Haak, W. *et al.* Ancient DNA from European Early Neolithic Farmers Reveals Their Near Eastern Affinities. *PLoS Biology* **8**, e1000536, doi:10.1371/journal.pbio.1000536 (2010).

10 Torroni, A., Achilli, A., Macaulay, V., Richards, M. & Bandelt, H. J. Harvesting the fruit of the human mtDNA tree. *Trends Genet* **22**, 339-345 (2006).

11 Cunliffe, B. & Koch, J. T. in *Celtic from the West: Alternative Perspectives from Archaeology, Genetics, Language and Literature* 384 (Oxbow Books, Oxford, 2010).

12 Pereira, L. *et al.* High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. *Genome Res* **15**, 19-24, doi:15/1/19 10.1101/gr.3182305 (2005).

13 Richards, M. B., Macaulay, V. A., Bandelt, H. J. & Sykes, B. C. Phylogeography of mitochondrial DNA in western Europe. *Ann Hum Genet* **62**, 241-260, doi:10.1046/j.1469-1809.1998.6230241.x (1998).

14 Roostalu, U. *et al.* Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: The near eastern and Caucasian perspective. *Molecular Biology and Evolution* **24**, 436-448 (2007).

15 Loogväli, E. L. *et al.* Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. *Mol.Biol.Evol.* **21**, 2012-2021 (2004).

16 Behar, D. M. *et al.* The Basque Paradigm: Genetic Evidence of a Maternal Continuity in the Franco-Cantabrian Region since Pre-Neolithic Times. *American Journal of Human Genetics* **90**, 486-493, doi:10.1016/j.ajhg.2012.01.002 (2012).

17 Behar, D. M. *et al.* A "Copernican" Reassessment of the Human Mitochondrial DNA Tree from its Root. *American Journal of Human Genetics* **90**, 1-10, doi:10.1016/j.ajhg.2012.03.002 (2012).

18 van Oven, M. & Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human Mutation* **30**, E386-394, doi:10.1002/humu.20921 (2009).

19 Green, R. E. *et al.* Analysis of one million base pairs of Neanderthal DNA. *Nature* **444**, 330-336, doi:10.1038/nature05336 (2006).

20 Hartmann, A. *et al.* Validation of Microarray-Based Resequencing of 93 Worldwide Mitochondrial Genomes. *Human Mutation* **30**, 115-122, doi:Doi 10.1002/Humu.20816 (2009).

21 Korlach, J. *et al.* Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol* **472**, 431-455, doi:10.1016/S0076-6879(10)72001-2 (2010).

22 Fu, Q., Rudan, P., Pääbo, S. & Krause, J. Complete Mitochondrial Genomes Reveal Neolithic Expansion into Europe. *PLoS ONE* **7**, e32473. doi:32410.31371/journal.pone.0032473 (2012).

23 Livi-Bacci, M. *A Consice History of World Population*. Fourth edn, (Blackwell Publishing, 2007).

24 Zimmermann, A., Hilpert, J. & Wendt, K. P. Estimations of population density for selected periods between the Neolithic and AD 1800. *Human Biology* **81**, 357-380, doi:10.3378/027.081.0313 (2009).

25 Ho, S. Y., Shapiro, B., Phillips, M. J., Cooper, A. & Drummond, A. J. Evidence for time dependency of molecular rate estimates. *Systematic biology* **56**, 515-522, doi:10.1080/10635150701435401 (2007).

26 Soares, P. *et al.* Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet* **84**, 740-759, doi:S0002-9297(09)00163-3-10.1016/j.ajhg.2009.05.001 (2009).

27 Mishmar, D. *et al.* Natural selection shaped regional mtDNA variation in humans. *P Natl Acad Sci USA* **100**, 171-176, doi:10.1073/pnas.0136972100 (2003).

28 Lacan, M. *et al.* Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. *P Natl Acad Sci USA* **108**, 9788-9791, doi:10.1073/pnas.1100723108 (2011).

29 Gamba, C. *et al.* Ancient DNA from an Early Neolithic Iberian population supports a pioneer colonization by first farmers. *Molecular Ecology* **21**, 45-56, doi:10.1111/j.1365-294X.2011.05361.x (2012).

30 Renfrew, C. *Archaeology and language : the puzzle of Indo-European origins*. Pimlico edn, (Pimlico, 1998).

Acknowledgements

We are indebted to Matt Kaplan and Ryan Spriggs at Arizona Research Laboratories, Division of Biotechnology, University of Arizona Genetics Core Facility, http://uagc.arl.arizona.edu/, Tyson Clark, Michael Brown, Kristi Spittle, and Matthew Boitano (Pacific Biosciences) for sequencing work, Jeremy Timmis for help with sonication protocols, and Clio Der Sarkissian, Robin Skeates, and Hubert Steiner for additional samples and context information. We thank the Australian Research Council (grant LP0882622), the Deutsche Forschungsgemeinschaft (Al 287/7-1 and Me 3245/1-1) and National Geographic's Genographic Project for funding.

Supplementary Information

Members of The Genographic Consortium

Syama Adhikarla, ArunKumar GaneshPrasad, Ramasamy Pitchappan & Arun Varatharajan Santhakumari, Madurai Kamaraj University, Madurai, Tamil Nadu, India; Elena Balanovska & Oleg Balanovsky, Research Centre for Medical Genetics, Russian Academy of Medical Sciences, Moscow, Russia; Jaume Bertranpetit, David Comas, Begoña Martínez-Cruz & Marta Melé, Universitat Pompeu Fabra, Barcelona, Spain; Andrew C. Clarke & Elizabeth A. Matisoo-Smith, University of Otago, Dunedin, New Zealand; Clio S. I. Der Sarkissian, University of Adelaide, South Australia, Australia; Matthew C. Dulik, Jill B. Gaieski, Amanda C. Owings, Theodore G. Schurr & Miguel G. Vilar, University of Pennsylvania, Philadelphia, Pennsylvania, United States; Angela Hobbs & Himla Soodyall, National Health Laboratory Service, Johannesburg, South Africa; Asif Javed, Laxmi Parida, Daniel E. Platt & Ajay K. Royyuru, IBM, Yorktown Heights, New York, United States; Li Jin & Shilin Li, Fudan University, Shanghai, China; Matthew E. Kaplan & Nirav C. Merchant, University of Arizona, Tucson, Arizona, United States; R. John Mitchell, La Trobe University, Melbourne, Victoria, Australia; Lluis Quintana-Murci, Institut Pasteur, Paris, France; Colin Renfrew, University of Cambridge, Cambridge, United Kingdom; Daniela R. Lacerda & Fabrício R. Santos, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; David F. Soria Hernanz & R. Spencer Wells, National Geographic Society, Washington, District of Columbia, United States; Pandikumar Swamikrishnan, IBM, Somers, New York, United States; Chris Tyler-Smith, The Wellcome Trust Sanger Institute, Hinxton, United Kingdom; Pedro Paulo Vieira, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; Janet S. Ziegle, Applied Biosystems, Foster City, California, United States.

1. Ancient DNA work

All ancient hg H individuals in this study were selected from a large pool of Neolithic samples (n=>400) from the Mittelelbe-Saale region in Saxony-Anhalt, Germany. This sample collection forms the core of an interdisciplinary, multi-centre project lead by the State Office for Heritage Management and Archaeology Saxony-Anhalt / State Museum for Prehistory Halle and the Bioarchaeometry group of the Johannes Gutenberg University of Mainz, Germany, and includes the Australian Centre of Ancient DNA (ACAD). All samples reported in this study are from recent excavations (2000 and younger), except for samples from the site of Derenburg, which were excavated from 1996-1999). A minimum of two samples per individual were collected under 'DNA-free' conditions and/or largely in situ following established protocols in collaboration with staff from the State Office in Halle[1]. Samples were not washed or treated after excavation and were kept refrigerated and/or in cooled conditions.

The majority of the ancient DNA (aDNA) work for this study was carried out at the specialised facilities of the Australian Centre of Ancient DNA (ACAD) following appropriate criteria to prevent/minimise contamination with modern DNA. DNA extractions, and sequencing of the mitochondrial control region HVS I samples DEB9 and DEB21 were carried out at the aDNA facilities of the Johannes Gutenberg University of Mainz, Germany. In addition, sample preparation, HVS-I sequencing and coding region SNP-typing of samples KAR6, KAR11, KAR 16, SALZ 18, SALZ 21, SALZ 57, SALZ77, EUL 41 and EUL 57 were also carried out in Mainz.

1.1. DNA extractions

A silica suspension was prepared by adding 6g of silicon dioxide (Sigma-Aldrich) to 50mL of DNA-free distilled water. The suspension was left for one hour to pellet larger grain sizes, before 40mL of the supernatant containing the finest particles was transferred to a new tube and kept overnight for further settling. Finally, a working silica suspension was created by discarding ~30mL of supernatant, retaining 10mL of the medium sized silica particles.

Preparation of tooth and bone samples for DNA extraction was carried out as previously described[2]. In previous studies we routinely used a phenol-chloroform based DNA extraction protocol that involved washing and concentration steps on Amicon filter units with a molecular weight cut-off of 30 and/or 50kDa, which results in a gradual loss of double-stranded DNA smaller than ~125bp (Amicon Ultra 4, User guide 2011). To recover DNA fragments of all sizes and especially from shorter fragments <100bp, we designed a customised DNA extraction protocol based on a standard silica-based extraction. On average

0.2g of tooth/bone powder were incubated overnight under constant rotation at 37°C in 4.44mL of lysis buffer consisting of 0.5M EDTA, pH 8.0; 0.5% N-lauroylsarcosine; 0.25mg/µL proteinase K. After lysis, the samples were centrifuged at 4,600 rpm for 1min and the supernatant transferred to a new 50mL tube. 125uL of medium-sized silica suspension (see above) and 16mL of in-house binding buffer (13.5mL QG buffer (Qiagen), 1X Triton, 20mM NaCl, 0.2M acetic acid (all Sigma-Aldrich)) were added and DNA was left to bind to silica overnight at room temperature under slow and constant rotation. The pH indicator included in the QG buffer provided the optimal pH conditions necessary for the binding of DNA to silica. On the third day the sample was centrifuged for 1min to pellet the silica particles and the supernatant was poured off. The pellet was transferred to a 1.5mL tube and washed three times by resuspension in 1mL 80% ethanol, centrifuged for 1min at 13,000rpm and the supernatant removed. The pellet was left to dry for 30min and subsequently resuspended in 200µL of pre-warmed (to 50°C) TE buffer (10mM Tris, 1mM EDTA) and incubated for 10min. After pelleting for 1min at 13,000rpm the supernatant was collected, aliquoted and stored at -18°C until further use.

1.2. PCR amplifications, HVS I sequencing and coding region SNP typing

All ancient hg H individuals in this study were selected from a large pool of Neolithic samples (n=>400) that were genotyped by direct sequencing of the mitochondrial hypervariable segment I (HVS I) and minisequencing of 22 coding region SNPs using a multiplex approach.

DNA was extracted from two independent samples for each individual. HVS I was amplified using a minimum of four short overlapping primer pairs, following established protocols as described previously[2,3]. Multiplex SNP typing of 22 haplogroup informative SNPs (GenoCoRe22) was carried out using a SNaPshot based protocol as described previously[3]. The GenoCoRe22 multiplex typing approach provided an ideal monitoring system for contamination of the ancient samples and non-template controls, as the PCR multiplex directly targets SNPs of all the major Eurasian haplogroups likely to constitute potentially contaminating lineages. Mitochondrial results were considered genuine and authentic when all sequences from replicated overlapping PCRs (a minimum of 6-8 fragments) produced unambiguous results in accordance with the GenoCoRe22 multiplex typing results from two independent extractions. One extract from each successfully typed individual was subsequently used for DNA library preparation. We also designed primer pairs targeting selected single nucleotide variants (discordant calls) from various quality score threshold settings as well as known characteristic sub-haplogroup SNPs (e.g. H1, H23, etc.) in

order to confirm or exclude these via direct sequencing from the original and/or independent extract (Supplementary Table 1).

In addition, the complete mt genomes of all staff at ACAD involved directly in the handling of the samples (P.B., W.H., C.J.A., and J.T.) and downstream steps of this study were sequenced to monitor potential contamination (Supplementary Table 2). DNA was extracted from swab samples and directly sequenced using standard protocols routinely used at the University of Arizona Genetics Core (UAGC)[4].

Supplementary Table 1. Primer sequences for additional coding region SNP confirmation via direct PCR and Sanger sequencing.

SNP	Name	Sequence 5' to 3'	PCR target
3010 (H1)	L0298	CAACAATAGGGTTTACGACCTC	71
	H0301	AACGAACCTTTAATAGCGGCTG	
10211 (H23)	L1018	TTACGAGTGCGGCTTCGAC	80
	H1021	AGAAGGTAATAGCTACTAAGAAGAATTTTA	
11152 (H26)	L1114	AACCACACTTATCCCCACCTT	78
	H1118	AAGTATGTGCCTGCGTTCA	
14060 (discordant)	L1403	CCTGACTAGAAAAGCTATTACCTAAAACA	80
	H1406	GCCTTTTTGGGTTGAGGTGAT	
14063 (discordant)	L1406	ACCAAATCTCCACCTCCATCA	79
	H1409	AGTGGGAAGAAGAAAGAGAGGAA	
2772 (H46)	L0277	AATGCAAACAGTACCTAACAAACC	72
	H0279	CGCCCCAACCGAAATTTTTAATG	
11893 (private)	L1186	TCGCTAACCTCGCCTTAC	66
	H1189	ACGTGGTTACTAGCACAGAGA	
10675 (discordant)	L1067	TTGCCATACTAGTCTTTGCCG	76
	H1070	CCATATGTGTTGGAGATTGAGACT	
10521 (discordant)	L1052	CTAGCATTTACCATCTCACTTCTA	67
	H1054	ATAGTAGGGAGGATATGAGGTG	

Supplementary Table S2. Haplotypes and polymorphic sites in mt genomes from staff members.

Staff	Haplotype	Differences to rCRS[5]
ADL3	H1a4	73, 263, (309.1C, 315.1C), 750, 1438, 3010, 4769, 8860,
ADL4	H1be	263, (309.2C, 315.1C), 750, 1438, 3010, 4769, 8860, 10750,
ADL5	H56c	263, (309.2C, 315.1C), 750, 1438, 4769, 8860, 11788,
ADL6	H1z1	263, (309.2C, 315.1C), 327, 750, 1438, 3010, 4769, 8860,

1.3. Ancient DNA Library preparation

DNA polishing/phosphorylation reactions were performed at 100µl final volume with 5 to 25µl of aDNA extract added to reactions comprising 50mM Tris-HCl pH 7.5, 10mM $MgCl_2$, 1mM ATP, 10mM Dithiothreitol, 250µg/ml rabbit serum albumin (RSA; Sigma), 400µM of each dNTP (Invitrogen), 50U T4 Polynucleotide Kinase (New England Biolabs, NEB), 10U DNA Polymerase I, Large (Klenow) Fragment (NEB), and 15U T4 DNA Polymerase (NEB). Thermocycling profiles consisted of 25°C for 15 min, 37°C for 15min, and 12°C for 15min. At 12°C, 10µl of 0.5M EDTA pH8.0 (Sigma) was added, followed by 550µl Qiagen Buffer PB1. DNA was purified using MinElute spin columns (Qiagen) as per the manufacturer's instructions.

Adaptor ligation reactions were performed at 60µl final volume with reactions comprising 62.8mM Tris-HCl pH7.6, 10mM $MgCl_2$, 1mM ATP, 2.8mM Dithiothreitol, 6% Polyethylene Glycol (PEG 6000), 2µM Adaptor UniHyb-A, 2µM Adaptor UniHyb-B, and 4,000U T4 DNA Ligase (NEB). The thermocycling profiles consisted of 20 cycles of 24°C

for 1min, 16°C for 30sec, and 8°C for 30sec. Then 300µl Qiagen Buffer PB1 was added to each reaction. DNA was isolated from the rest of the reaction components using MinElute spin columns (Qiagen) as per the manufacturer's instructions. The partially double stranded adaptor UniHyb-A comprised the oligonucleotides UniHyb-Af GGTGTTGTTAGGAATGCGAGA and UniHyb-Ar TCTCGCATTCCTAA. The partially double stranded adaptor UniHyb-B comprised the oligonucleotides UniHyb-Bf AGGATAGGTCGTTGCTGTGTA and UniHyb-Br TACACAGCAACGA. UniHyb-A and UniHyb-B were formed by hybridisation with a thermocycling profile of 95°C for 2min, then 75°C for 20sec, followed by a ramp from 75°C to 10°C at 2°C/min increments.

Polymerase 'fill-in' reactions, to remove nicks and to create fully double-stranded adaptor-tagged aDNA, were performed at 30µl final volume with reactions comprising 20mM Tris-HCl pH 8.8, 10mM $(NH_4)_2SO_4$, 10mM KCl, 2mM $MgSO_4$, 0.1% Triton X-100, 250µM of each dNTP, and 16U Bst DNA Polymerase, Large Fragment (NEB). The thermocycling profile was 37°C for 30min. Then 150µl Qiagen Buffer PB1 was added to each reaction. DNA was purified using MinElute spin columns (Qiagen) and eluted into 21µl as per the manufacturer's instructions.

PCR amplification reactions, to create 'master' DNA libraries, were performed in 3 x 44µl volumes per original sample/extract, with 7µl of eluted DNA added per tube to reactions comprising 1x AmpliTaq Gold buffer II, 2.5mM $MgCl_2$, 2.5U AmpliTaq Gold (Applied Biosystems), 250µM of each dNTP (Invitrogen), and 0.5µM of PCR primers UniHyb-PCR-A (GGTGTTGTTAGGAATGCGAGA) and UniHyb-PCR-B (AGGATAGGTCGTTGCTGTGTA), see also Supplementary Figure 1. The thermocycling profile consisted of 94 °C for 11 min, followed by 12 cycles of 30sec at 95°C, 30sec at 60°C and 1min (+2 sec/cycle) at 72°C, followed by a final 10min at 72°C. The 3 x 44µl volume reactions were pooled. From each original sample/extract 'master' DNA library pool, 2.5µl was then added to 4 x 35µl reactions comprising: 1X AmpliTaq Gold buffer II, 2.5mM $MgCl_2$, 250µM of each dNTP, 1.0U AmpliTaq Gold (Applied Biosystems); and 0.5µM of PCR primers UniHyb-PCR-A and UniHyb-PCR-B. The remainder of the 'master' DNA libraries were archived at -80°C. Once reamplified using the above thermal profile and pooled, these 4 x 35µl amplification reactions were purified using MinElute spin columns (Qiagen) and eluted into 15µl as per the manufacturer's instructions. These comprised the 'primary' DNA libraries and amplification products were sized and quantified via gel electrophoresis against quantified size markers (HyperLadder™ V, Bioline) and a Nanodrop 2000 (Thermo Scientific).

1.4. Hybridisation-based enrichment of human mtDNA target sequences within libraries

The basic design for hybridisation of tracer (in this case library) DNA and biotinylated driver (in this case human mitochondrial probe) DNA sequences was previously described by Patel and Sive (2001)[6]. The main differences were as follows. First, in addition to the described HEPES/NaCl-based hybridisation conditions, SSC-based hybridisation conditions were also used with Na^+ concentrations between 660-990mM. Second, hybridisation reactions were carried out in a total volume of 30µl. Third, 300-400ng of MinElute-concentrated 'primary' library DNA and 85-120ng of biotinylated probe DNA were used. Fourth, the thermocycling profiles used for hybridisation were 95°C for 5min, followed by 14-18h at 50 or 55°C. Fifth, and finally, the two library primers UniHyb-PCR-A and UniHyb-PCR-B were included as part of the hybridisation mix at 0.67-1.0µM.

There were two reasons behind the inclusion of the library amplification PCR primers in the hybridisation mix. First, by annealing to their complementary sequences in the adaptors flanking the ancient DNA insert sequences in the PCR amplified libraries, UniHyb-PCR-A and UniHyb-PCR-B could act as 'blocking oligos'[7] to minimise unwanted hybridisation between the adaptor-tagged flanking regions of otherwise unrelated single-stranded library DNA molecules (Supplementary Figure 1). Second, following binding of biotinylated probe molecules to Streptavidin beads and stringency washes to remove non- or weakly-hybridised single-stranded library DNA molecules, incubation with dNTPs and a DNA polymerase with strand-displacing activity would allow primer extension to neatly disrupt the double-stranded region of stable hybridisation between human mitochondrial probe DNA sequences and single-stranded library DNA molecules that had inserts with complementary sequences. The net effect of this polymerase 'filling-in' reaction would be to free the originally single-stranded library DNA molecules captured by the probe molecules by hybridisation: the biotinylated probe molecules would remain bound to the magnetic Streptavidin beads and the discrete now-'filled-in' double-stranded library DNA molecules could be removed cleanly in the free supernatant using a magnetic rack (Invitrogen) which would keep the bead-bound probe molecules pelleted adjacent to the magnet (Supplementary Figure 1).

Following the 14 to 18h overnight hybridisation step, the biotinylated probe molecules in the 30µl hybridisation reactions were bound and immobilised to magnetic Streptavidin beads as previously described[8]. Successive stringency washes, to remove non- or weakly-hybridised single-stranded library DNA molecules, used decreasing salt and increasing temperature and followed the following profile: 2X SSC/0.1% SDS at 37°C for 1min; 2X SSC/0.1% SDS at 42°C for 10min; 1X SSC/0.1% SDS at 43°C for 10min; 0.5X SSC/0.1% SDS at 44°C for 10min; 0.5X SSC/0.1% SDS at 45°C for 10min. At this point, the beads

(with attached biotinylated probe molecules and stably hybridised single-stranded library DNA molecules with complementary insert sequences) were attached to the tube side via the magnet and the supernatant was removed via pipetting and discarded. The beads were then immediately resuspended in buffer conditions to suit a strand-displacing polymerase activity.

Two strand-displacing polymerases, the Klenow Fragment (3′→5′ exo⁻) and the *Bst* DNA Polymerase, Large Fragment (NEB) were used. With the Klenow Fragment (3′→5′ exo⁻), the reactions were performed at 60µl final volume comprising 15mM Tris-HCl pH 8.0, 50mM KCl, 5.5mM $MgCl_2$, 250µM of each dNTP, and 5U Klenow Fragment (3′→5′ exo⁻). The tubes were incubated at 37°C for 15-30min with regular agitation to keep the beads in suspension. The reaction tube was then applied to the magnetic rack and the 60µl supernatant transferred to a fresh 200µl tube. This tube was immediately incubated at 80°C for 20min to inactivate the enzyme. With the *Bst* DNA Polymerase, Large Fragment, the reactions were performed at 35µl final volume comprising 15mM Tris-HCl pH 8.0, 50mM KCl, 10.0mM $MgCl_2$, 200µM of each dNTP, and 100µg/ml of BSA. The beads were resuspended in the buffer – minus the *Bst* DNA Polymerase, Large Fragment enzyme – and pre-heated to 60°C. The enzyme was added and the tubes were incubated at 60°C for 5min with regular agitation to keep the beads in suspension. The reaction tube was then applied to the magnetic rack at 60°C and the 35µl supernatant transferred to a fresh 200µl tube. This tube was immediately incubated at 80°C for 20min to inactivate the enzyme.

Using either approach, the heat-inactivated supernatant was split between several (generally 4-8) PCR reamplification reactions (total combined volume 140µl), designed so that upon the addition of the sub-portion of Klenow or *Bst* buffer, the final composition of the reactions would be 1X AmpliTaq Gold buffer II, 2.5mM $MgCl_2$, 250µM of each dNTP, 1.0U AmpliTaq Gold (Applied Biosystems), and 0.5µM of PCR primers UniHyb-PCR-A and UniHyb-PCR-B using the above thermal profile. Amplification reactions were pooled and library amplificons purified using MinElute spin columns (Qiagen) and eluted into 15 µl as per the manufacturer's instructions. These comprised the 'first enrichment' DNA libraries and amplification products were sized and quantified via gel electrophoresis against size markers (HyperLadder™ V, Bioline) and a Nanodrop 2000 (Thermo Scientific). In general, the whole hybridisation / enrichment / reamplification cycle was repeated three times, to produce 'third enrichment' DNA libraries highly enriched for short fragments of endogenous mtDNA sequence within the ancient DNA libraries.

The above system underwent development and optimisation over the period of this research. A number of parameters were varied in order to optimise the enrichment and coverage of human mtDNA sequences within libraries. Overall, a number of observations can be made about the above hybridisation / enrichment / reamplification steps. Most importantly,

of all the parameters varied, the only ones that appear to have made a substantial difference to the experimental outcome were the stringency wash steps. High temperatures and/or low salt concentrations generally led to poor subsequent library reamplification (presumably as many of the short human mtDNA inserts within libraries – generally 20-70 bases – did not hybridise to probe DNA efficiently beyond a certain point). Lower temperatures / higher salt concentration led to poorer coverage of the mt genome following enrichment, presumably due to the retention of significantly higher proportions of non-mtDNA sequences during lower stringency washes (data not shown). Empirically, the stringency wash conditions described above seemed to work well with DNA extracted and made into libraries from generally well-preserved archaeological remains. Switching between HEPES/NaCl-based versus SSC-based hybridisation conditions, using between 300-400ng of MinElute-concentrated 'primary' library DNA, using between 85-120ng of biotinylated probe, overnight hybridisation for 14 to 18h at either 50 or 55°C, or using between 0.67-1.0μM UniHyb-PCR-A and UniHyb-PCR-B primers all appeared to make little or no discernable difference.

1.5. Preparation of biotinylated human mitochondrial DNA probes

Total genomic DNA from a present-day individual belonging to mitochondrial hg J1c8 (with SNP differences from the revised Cambridge Reference Sequence (rCRS)[5] at nucleotide positions: 73G, 185A, 228A, 295, 462, 489, 2706, 3010, 4216, 7028, 10084, 10398, 11251, 11719, 12612, 13708, 14766, 14798, 15452A, 16069, 16126, 16261, 16265 and 16319) was extracted from a buccal swab using the DNeasy Blood & Tissue Extraction kit (Qiagen). This hg J1c8 mitochondrial genome was amplified in two overlapping fragments by long-range PCR (LR-PCR) using the PCR primer pair combinations L06363/H14799 (8476bp) and L14759/H06378 (8340bp). Primer details are: L06363 (ACCATCTTCTCCTTACACCTAGCAG), H14799 (GGTGGGGAGGTCGATGA), L14759 (AGAACACCAATGACCCCAATAC) and H06378 (GATGAAATTGATGGCCCCTAA). Probe was generated in two ways: (a) via LR-PCR with Biotin-dNTPs included in the dNTP mix, followed by DNA fragmentation; (b) via LR-PCR without Biotin-dNTPs, followed by DNA fragmentation then 3'-biotinylation with Biotin-ddUTP using the enzyme Terminal Transferase (TdT).

(a) Biotinylated probe DNA was generated by LR-PCR (Expand LR dNTPack, Roche) according to the manufacturer's recommendations. Biotin-dNTPs (NEB Biotinylated dNTP Mixture) were incorporated into the PCR reaction at a ratio of 1:10 with regular dNTPs. The whole mitochondrial genome was amplified in eight separate 25μl LR-PCR reactions using the two primer sets L06363/H14799 (8476bp) and L14759/H06378 (8340bp). In total 42 cycles of PCR were carried out using the Bio-Rad Tetrad 2 Thermal Cycler.

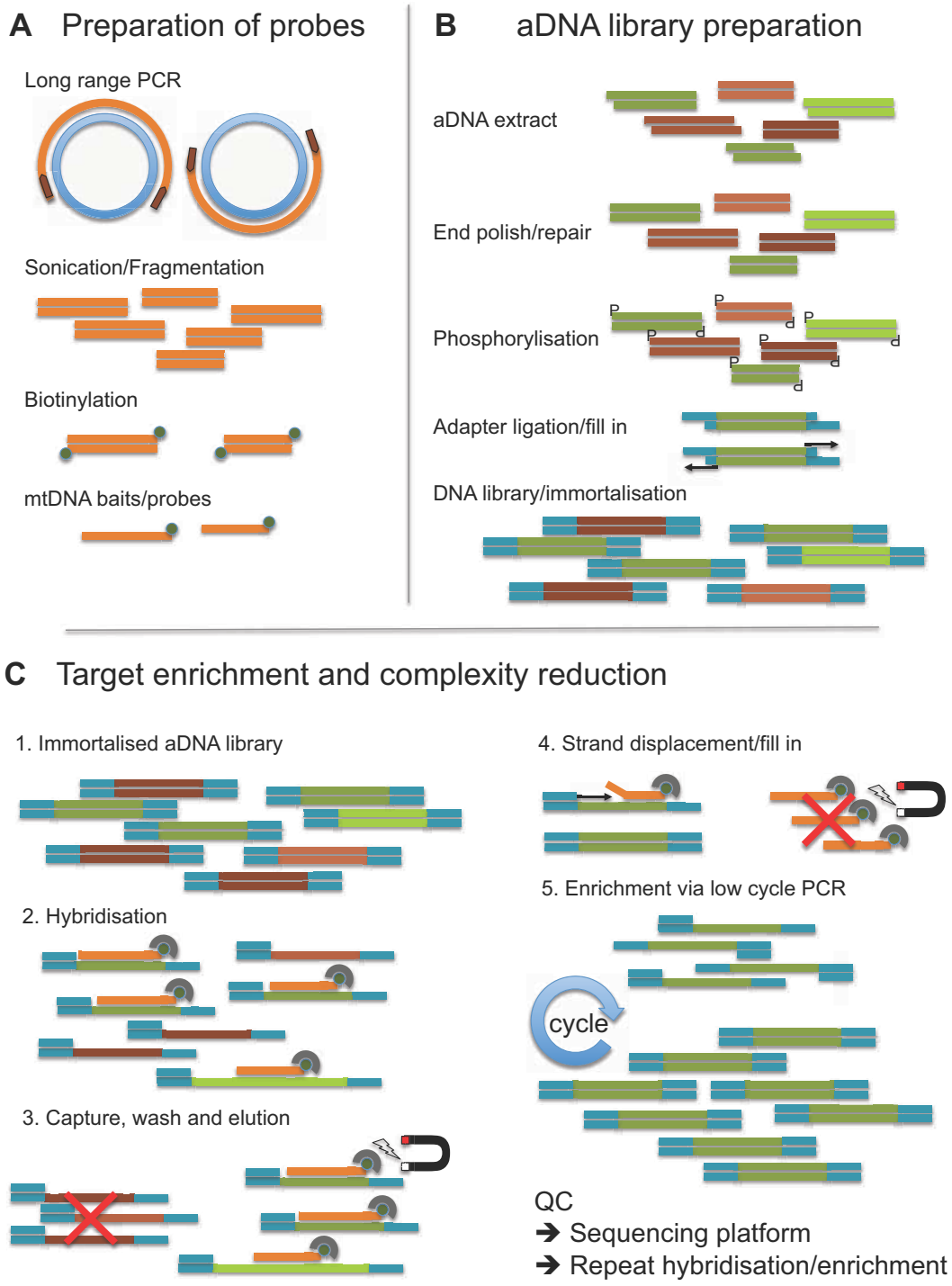
Cycling conditions included: initial denaturation at 92°C for 2mins; 10 cycles of 92°C for 10sec, 60°C for 15sec, 68°C for 8min 40sec; followed by 32 cycles of 92°C for 10sec, 60°C for 15sec, 68°C for 8 min 40sec – increasing by 15 sec/cycle; final extension at 68° for 10mins; followed by a hold at 4°C. PCR reactions were pooled for each product and loaded onto 1% Agarose gels. Bands were visualised by ethidium bromide staining prior to excision of the correct-sized band under UV light with a clean, sharp scalpel blade. Purification of bands was carried out using the QIAquick Gel Extraction Kit (Qiagen) and eluted in 30µl as per manufacturer's instructions. Quantification was performed using gel electrophoresis against quantified size markers (HyperLadder™ V, Bioline) and a Nanodrop 2000 (Thermo Scientific). DNA was diluted to 120µl total volume with distilled water as a requirement for the sonicator prior to shearing. Fragmentation of DNA to the desired size range was achieved by discontinuous sonication (Microtip sonicator, Thomas Optical & Co. Pty. Ltd) at high speed, amplitude 6, for 4 x 1min intervals, no ice. Sonication was performed in time intervals to avoid successive heat build-up that can damage the DNA and the sample was subjected to centrifugation during each time interval, using a benchtop centrifuge. Post sonication, DNA was purified and concentrated using the QIAquick PCR Purification Kit (Qiagen) and eluted in 30µl as per the manufacturer's instructions. The resulting biotinylated probe DNA was analysed on a 1% Agarose gel, with the resulting smear in the size range 200-600bp. DNA concentration was estimated using the Nanodrop. DNA from both halves of the mitochondrial genome was pooled to produce an equimolar concentration of biotinylated probe DNA across the genome, ready for hybridisation.

(b) Biotinylated probe DNA was generated by LR-PCR (Expand LR dNTPack, Roche) according to the manufacturer's recommended conditions. The whole mitochondrial genome was amplified in eight separate 25ul LR-PCR reactions using the two primer sets L06363/H14799 (8476bp) and L14759/H06378 (8340bp). In total 42 cycles of PCR were carried out using the Bio-Rad Tetrad 2 Thermal Cycler. Cycling conditions included: initial denaturation at 92° for 2 mins; 10 cycles of 92°C for 10sec, 60°C for 15sec, 68°C for 8min 40sec; followed by 32 cycles of 92°C for 10sec, 60°C for 15sec, 68°C for 8min 40sec – increasing by 15sec/cycle; final extension at 68° for 10mins; followed by a hold at 4°C. PCR reactions were pooled for each product and loaded onto 1% Agarose gels. Bands were visualised by ethidium bromide staining prior to excision of the correct-sized band under UV light with a clean, sharp scalpel blade. Purification of bands was carried out using the QIAquick Gel Extraction Kit (Qiagen) and eluted in 30 ul as per the manufacturer's instructions. Quantification was performed using a Nanodrop 2000 (Thermo Scientific). DNA was diluted to 120ul total volume with distilled water as a requirement for the sonicator prior

to shearing. Fragmentation of DNA to the desired size range was achieved by sonication (Microtip sonicator, Thomas Optical & Co. Pty. Ltd) at high speed, amplitude 6, for 4min, no ice. Post sonication, DNA was purified and concentrated using the QIAquick PCR Purification Kit (Qiagen) and eluted in 30ul as per the manufacturer's instructions. The resulting DNA was analysed on a 1% Agarose gel, with the resulting smear in the size range 200-600bp. DNA concentration was estimated using gel electrophoresis against quantified size markers (HyperLadder™ V, Bioline) and a Nanodrop 2000 (Thermo Scientific). DNA from both halves of the mitochondrial genome was pooled to produce an equimolar concentration of biotinylated probe DNA across the genome, ready for biotinylation via Biotin-16-ddUTP and the enzyme Terminal Transferase (TdT). Prior to 3' end-labelling, sonicated probe DNA was made single stranded by heating at 95°C for 5min, then immediately placed on wet ice for 5min. Reactions for the 3' end-labelling of probe DNA were performed at 50µl final volumes comprising: 50mM Potassium Acetate; 20mM Tris-Acetate pH 7.9; 10mM Magnesium Acetate; 0.25mM Cobalt Chloride; sonicated mitochondrial probe DNA at 10pmol of 3' ssDNA ends; 0.1mM Biotin-16-ddUTP (Enzo); 40U Terminal Transferase enzyme (NEB). The thermocycling profiles were: 37°C for 60min; then 70°C for 10min. DNA was isolated from the rest of the reaction components using Qiagen Nucleotide Removal Kit spin columns and eluted into 30ul as per the manufacturer's instructions.

Considerations of probe biotin labelling and library elution via strand-displacing polymerases: as described above, there were two versions of probe synthesis and preparation: (a) via LR-PCR with Biotin-dNTPs included in the dNTP mix, followed by DNA fragmentation; and (b) via LR-PCR without Biotin-dNTPs, followed by DNA fragmentation then 3'-biotinylation with Biotin-16-ddUTP using the enzyme Terminal Transferase (TdT). In addition, two strand-displacing polymerases were used to disrupt the double-stranded region of stable hybridisation between human mitochondrial probe DNA sequences and single-stranded library DNA molecules that had inserts with complementary sequences, and thereby elute library DNA molecules: (a) the Klenow Fragment (3′→5′ exo$^-$); and (b) the *Bst* DNA Polymerase, Large Fragment. Initially, (a) and (a) were used together. Although this approach worked extremely well, on one occasion (probe batch October 2010) probe sequences could be identified in enriched library DNA sequences (almost exclusively in the overlapping region between the probe amplicons). Presumably, this occurred due to interactions between library molecules with mtDNA inserts, library oligos and free 3'-ends of probe molecules. Fortunately, we were able to eliminate these occasional hg J1c8 intrusions via direct PCR on independent ancient DNA extracts from the same archaeological individuals. However, this encouraged the development of the combined (b) and (b) approach. The rationale was that 3'-

Biotin-16-ddUTP labelling the probe would prevent probe molecules from extending under any circumstances. In addition, using the *Bst* DNA Polymerase, Large Fragment at 60 °C should also drastically reduce any potential spurious, non-specific, hybridisation of library primers to probe (compared to those that might occur at 37°C). Once this approach had been adopted, we didn't see a single instance of any identifiable probe SNP in any enriched libraries, for either human or non-human samples (data not shown).



Supplementary Figure 1. Schematic representation of working steps involved in a) preparation of probes, b) ancient DNA library preparation, and c) targeted enrichment of mitochondrial DNA.

2. Affymetrix Mitochip v2.0

2.1 Preparation of enriched libraries as MitoChip probes

Following the final (third) round of hybridisation / enrichment / reamplification, ~1/10 of the enriched library was reamplified with the dU-containing primers: U-UniHyb-Af (GGTGTTGUTAGGAAUGCGAGA) and U-UniHyb-Bf (AGGATAGGUCGTTGCUGTGTA). Following reamplification, the dU bases could subsequently be digested with a combination of Uracil DNA glycosylase (UDG) and the DNA glycosylase-lyase Endonuclease VIII. (NEB's USER Enzyme is a pre-mixed combination of these two enzymes.) UDG catalyses the excision of uracil bases, forming an abasic (apyrimidinic) site while leaving the phosphodiester backbone intact, whereas the lyase activity of Endonuclease VIII breaks the phosphodiester backbone at the 3′ and 5′ sides of the abasic site so that base-free deoxyribose is released. Effectively then, treatment of the U-UniHyb-Af / U-UniHyb-Bf amplified libraries with the USER Enzyme trims off all but 6 bases (A) or 5 bases (B) of the UniHyb-PCR-A and UniHyb-PCR-B library adaptors. This minimises any non-mtDNA hybridisation between unrelated single-stranded library DNA molecules via their adaptor sequences, maximising the opportunities for authentic enriched human mtDNA sequences to hybridise to the MitoChip. The specific PCR amplification conditions used were 1X AmpliTaq Gold buffer II, 2.5mM $MgCl_2$, 250µM of each dNTP, 1.0U AmpliTaq Gold (Applied Biosystems), and 0.5µM of PCR primers U-UniHyb-Af and U-UniHyb-Bf. The thermocycling profile was 94°C for 10min; followed by 25-30 cycles of 30sec at 95°C, 30sec at 60°C and 30sec (+1 sec/cycle) at 72°C; followed by a final 20min at 72°C. DNA was purified using MinElute spin columns (Qiagen) as per the manufacturer's instructions. USER Enzyme digestion took place for 2-3 hours at 37°C in 60µl final volume reactions comprising 15mM Tris-HCl pH 8.0, 50mM KCl, 1.5mM $MgCl_2$, 3U USER Enzyme. The entire reaction was then cleaned up by passing through a BioRad P-30 column as per the manufacturer's instructions and purified DNA sized and quantified as described above. The isolated DNA was then biotin-labelled as described below (Section 2.2).

2.2. Affymetrix Mitochip v2.0 background and sample preparation

The Affymetrix Mitochip v2.0 is an oligonucleotide tiling array for high-throughput resequencing analysis of the human mitochondrial (mt) genome[9]. For each nucleotide position (np) of the mt genome interrogated, the array possesses eight 25-mer probes - four each to match the heavy and the light strands, respectively, of the rCRS mt genome[5]. Each of these strand-specific 25-mers varies only at the central position. For each np, this allows all four possible alleles (A, T, C or G) to be interrogated on both strands within a tight local sequence context – via the hybridisation of fluorescently labelled fragments of sample DNA. The Mitochip v2.0 also carries additional local context probes to interrogate mtDNA sequence variants for 500 of the most common haplotypes (e.g. particular insertions, deletions, and adjacent or closely spaced SNPs) based on the FBI database[9]. There are a number of reasons why resequencing via oligonucleotide tiling arrays might be an attractive genotyping strategy for researchers dealing with samples that yield highly damaged and degraded DNA.

With some forensic, environmental, or formalin-fixed paraffin-embedded (FFPE) archived samples and medical biopsies, DNA degradation and physical fragmentation can be considerable and irreversible. Forensic, environmental, and in particular ancient/archaeological samples (including archived museum specimens covering unique and irreplaceable now-lost genetic diversity from both extant and extinct species), can contain minute amounts of highly damaged and degraded endogenous DNA. Worse still, up to 99% or more of DNA templates extracted from samples like these can be non-endogenous post-mortem and soil bacterial DNA sequences. DNA extracts with a high background complexity like these sometimes contain few or no endogenous DNA templates >100bp in length (as determined via PCR amplification with a primer pair targeting a 100bp region). However, a key finding in recent years has been that as one interrogates shorter and shorter DNA templates in ancient DNA extracts, the effective copy numbers of DNA targets in that size range increases exponentially. Evidence from a number of independent studies has conclusively demonstrated this negative exponential relationship between fragment size and template copy number[8,10-13].

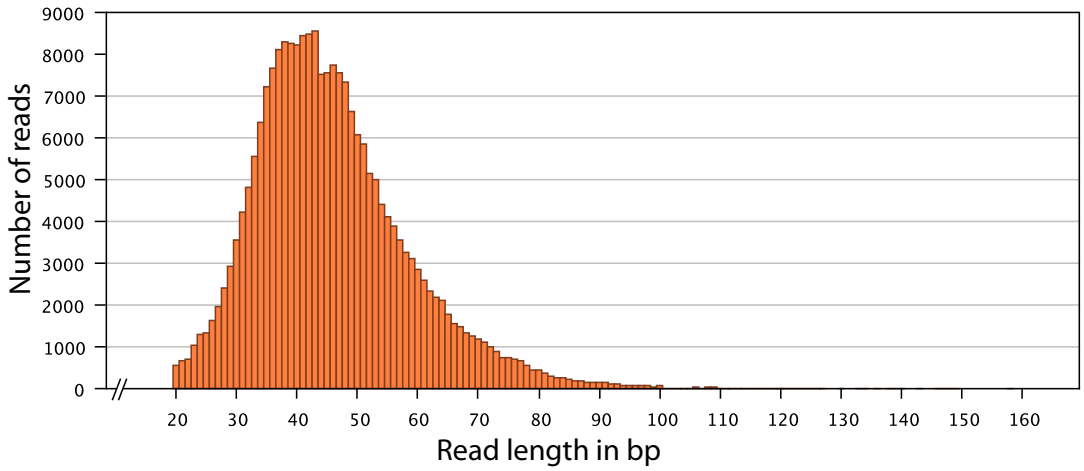
This relationship between aDNA fragment size and copy number provides the rationale behind the use of the Mitochip v2.0. Provided short fragments of endogenous human mtDNA sequences of ~20-25 bases and above could be significantly enriched (over and above fragments of non-mt human DNA and the high levels of non-endogenous bacterial DNA sequences typically intrinsic to archaeological samples), the Mitochip v2.0 could interrogate every nucleotide position (np) of the mt genome via its overlapping 25 base windows. The chip's ability to SNP-type fragments of aDNA of 20-25 bases and above would

maximise the chances of genotyping mt genomes from even highly marginal archaeological samples thousands of years old. Prior to the use of the Mitochip v2.0 as an aDNA genotyping tool, therefore, we had to develop and optimise new methodologies designed: (a) to efficiently extract short fragments of damaged and degraded DNA from archaeological remains; (b) to immortalise this extracted DNA as amplifiable libraries; (c) to enrich for human mtDNA target sequences within these libraries via hybridisation-based DNA-capture protocols; and (d) to label these mt-enriched DNA libraries for use as probes in Mitochip-based resequencing analyses of the human mt genome[9].

In everyday use, with DNA extracts from modern human samples, the 16569 bp mtDNA genome is typically amplified via two or three overlapping long-range (LR-)PCR targets[9,14-16]. Following LR-PCR, amplification products undergo fragmentation, labelling, and finally hybridisation, as described in the Affymetrix GeneChip CustomSeq Resequencing Array Protocol[17]. Clearly however, LR-PCR with target lengths over 5kb is impossible for highly damaged and degraded endogenous DNA extracted from marginal forensic or archaeological samples. For example, in the range 126-179 bp used to investigate the control region in short overlapping fragments[2], one of our ancient DNA extracts failed to consistently yield amplicons (individual HQU4, which nevertheless could be placed into hg H7d5 using the whole mtDNA Mitochip v2.0 approach described herein).

Both theoretical considerations and empirical evidence indicate two opposing influences on library insert sizes during our mtDNA enrichment process (Section 1.4.). Hybridisation during the targeted DNA-capture step tends towards more efficient enrichment of longer mtDNA insert sequences, due to the increased stability of longer annealed hybrids between library and probe sequences at any given hybridisation, stringency wash temperatures and salt concentrations. In contrast, post-hybridisation library PCR reamplification preferentially amplifies smaller amplicons (i.e. library constructs with shorter aDNA sequence fragments inserted between library adaptors)[18]. As shown previously, damaged and degraded ancient DNA fragments are (on average) shorter than non-ancient molecules, meaning adaptor-tagged ancient mtDNA fragments will generally be strongly preferentially amplified over any present-day / 'modern' adaptor-tagged contaminant mtDNA present[11,7].

Between these 'push-and-pull' influences on library insert size, our hybridisation-based DNA-capture and library (adaptor) PCR reamplification protocols allowed us to enrich human mtDNA target sequences largely in the 20-70 bp range. Over 95% were within this size range based on a total of 232,347 reads from Pacific Biosystems data from six libraries enriched for human mtDNA sequences. Here, only 10666 (4.6%) were longer than 70bp with a maximum read length reaching 158bp (Supplementary Figure 2).



Supplementary Figure 2. Distribution of number and length of human mtDNA reads from Pacific Biosciences data (232,347 reads in total).

The vast swathe of endogenous ancient human mtDNA fragments between ~20-to-50 bases (Supplementary Figure 2) targeted in this study would have been too short to be amplified directly by traditional PCR-based approaches (assuming typical forward and reverse PCR primers of ~25 bases). In contrast to traditional PCR-based approaches, therefore, hybridisation-based DNA-capture vastly expands both the potential aDNA template copy numbers available (via targeting <50 base aDNA molecules) and the mitochondrial genome coverage (via targeted enrichment using a whole mitochondrial genome probe).

Endogenous ancient mtDNA molecules largely within the 20-70bp range have no need for a fragmentation step prior to labelling as MitoChip v2.0 probe, as they are already ideally sized for a chip-based resequencing/genotyping approach based on hybridisation to the strand-specific 25-mers on the array. Assuming the exponential increase in copy number with decreasing target size continues, our approach therefore maximises the chances of assembling mt genome sequences from the most difficult, poorly preserved, samples[8,13]. However, DNA molecules fragmented to <20 bases are effectively beyond recovery and analysis, as fragments smaller than this cannot be captured efficiently by hybridisation or meaningfully tackled bioinformatically[19].

Following DNA extraction, library immortalisation, DNA-capture enrichment for human mtDNA target sequences within these libraries, enriched library reamplification, and pre-labelling preparation (Sections 1.3-1.5.), we therefore bypassed the fragmentation step in the Affymetrix CustomSeq Resequencing Array Protocol[17]. Instead we directly labelled prepared libraries using the following protocol:

1. 4-6 µg of enriched library DNA underwent biotin labelling using Terminal Deoxynucleotidyl Transferase (TDT) as per the Affymetrix GeneChip Whole Transcript Sense Target Labelling Assay Manual (P/N 701880, rev. 4).

2. Labelled samples were hybridised to Affymetrix GeneChip Human Mitochondrial Resequencing 2.0 Arrays for 17 hours at 49°C.

3. Arrays were washed, stained and scanned as per the GeneChip CustomSeq Resequencing Array Protocol (P/N 701231, rev. 5).

4. Affymetrix GeneChip Command Console software (v3.2) was used to generate CEL files, which were then analysed using GeneChip Sequence Analysis Software (GSEQ v4.1, Affymetrix).

An intrinsic aspect of resequencing arrays is an inevitable trade-off between call rate and accuracy of analyses[9,16,20]. Previous studies using both the Mitochip v2.0 and other resequencing microarrays have empirically optimised parameters. For example, after using quality score threshold settings (QST) of 2, 3, 6, 9, 12, and 30, Hartmann et al. (2008) settled on the haploid model with a QST of 3 and no reliability rule filter as the favoured parameters for their analysis of 93 worldwide (present-day) mitochondrial genomes[16]. However, although previous studies that used the Mitochip v2.0 and other resequencing microarrays are instructive and informative, there are at least two key aspects of our study that do not reflect any previous analyses by resequencing microarrays and which meant that we needed to empirically establish the most favourable parameters for ourselves.

First, in previous Mitochip studies 100% of the input probe DNA corresponded to human mtDNA sequences from the sample-of-interest However, in our case, even with successive rounds of targeted DNA-capture enrichment of human mtDNA sequences, followed by library re-amplification, the final enrichment libraries used to generate probe will inevitably be composed of <100% human mtDNA sequences. As described above, cloning and next-generation library sequencing of enriched DNA libraries suggests that 65-85% of input mt-enriched library probe onto chips is likely to correspond to ancient human mtDNA sequences. Second, as has been shown in many studies, the amplification of ancient DNA generates sequence changes due to miscoding lesion DNA damage[8,21-23]. This damage derived DNA sequence variation within the labelled Mitochip probe therefore added an additional layer of complexity.

2.3. Affymetrix Mitochip v2.0 validation for ancient DNA

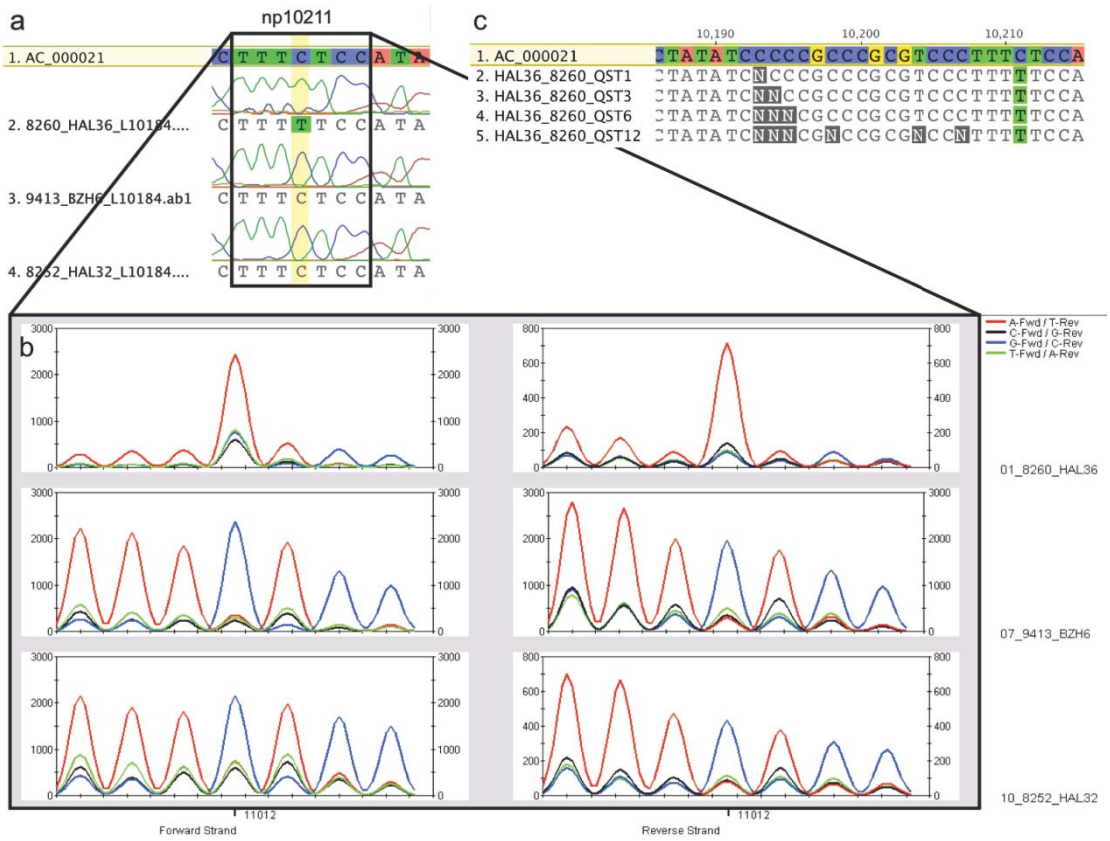
The performance of the Mitochip v2.0 on ancient DNA libraries enriched for human mtDNA sequences was assessed by several criteria.

First, for all 39 archaeological individuals, the original DNA extracts from which the initial master and primary libraries were prepared, as well as duplicate extracts from separate bone/tooth samples from the same archaeological individuals (for replication), were tested

independently over 413bp of HVSI via short overlapping PCR amplicons. In one case, ancient DNA extracts failed to yield every HVSI PCR amplicon (HQU4). In total, duplicate direct PCR and Sanger sequencing analyses of all 39 hg H individuals identified 16 SNP differences from the rCRS in HVSI (Supplementary Table 3). From the Mitochip runs using our mtDNA-enriched libraries as probes, 15 of these 16 SNPs were also identified for all levels, QST 1, 3, 6 and 12. One mutation (16213A in individual QUEVIII4) was called correctly for QST 1, but called as 'N' by QST 3, 6 and 12.

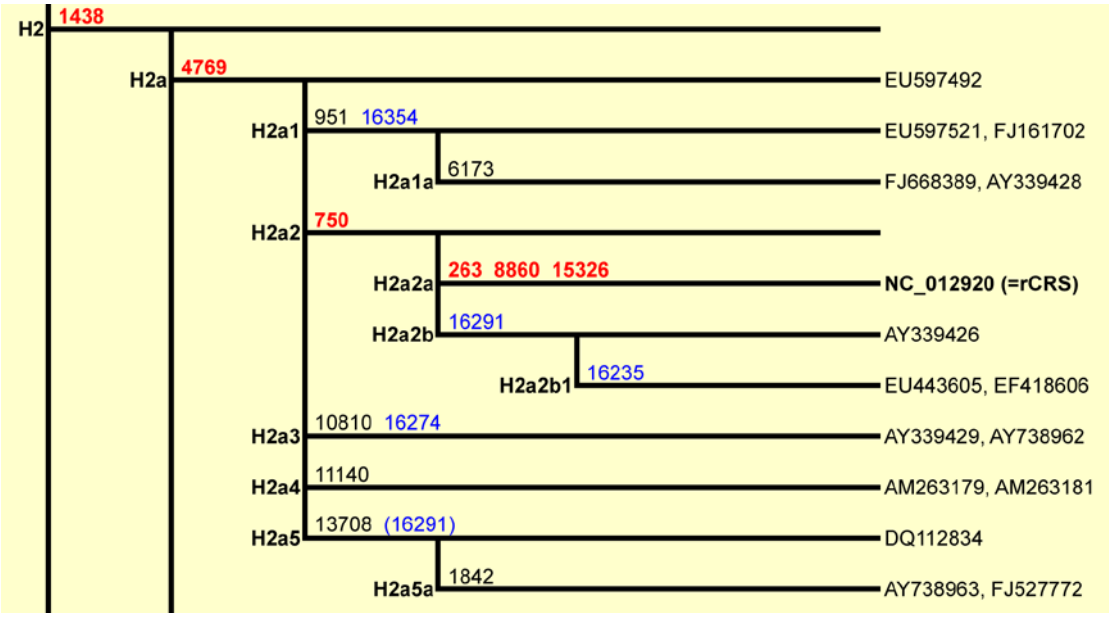
Second, 363 bp of HVSII were also amplified in a number of short overlapping fragments from 9 of the ancient DNA extracts and typed by Sanger sequencing as described previously[24]. Direct PCR and Sanger sequencing analyses of data from these individuals identified a further 13 SNP differences from the rCRS (ignoring unstable C-stretch insertions; Supplementary Table 3). From the Mitochip v2.0 runs using our mtDNA-enriched libraries as probes, 12 of these SNPs were also identified for all levels, QST 1, 3, 6 and 12. One mutation (00152C from SALZ 57A) was called correctly for QST 1, but called as 'N' by QST 3, 6 and 12.

Third, an additional six (previously described) subhaplogroup-defining SNPs identified via the Mitochip v2.0 were also independently confirmed (Supplementary Table 3). Thirty-five out of 159 SNPs (31%) were directly sequenced. All six SNPs were independently tested using newly designed primer pairs (Supplementary Table 1) via direct PCR from the extract used to make the original library (and/or the duplicate extract) followed by Sanger sequencing. All six confirmed the SNPs identified via the Mitochip v2.0 runs. This independent confirmation was particularly important for haplogroups defined by a single SNP and phylogenetic branching point, such as hg H23 (defined by 10211T, Supplementary Figure 3) and hg H26 (defined by 11152C) for the Early Neolithic LBK individuals HAL36 and HAL32, respectively.



Supplementary Figure 3. Examples of data from different sequencing platforms highlighting a diagnostic SNP at np 10211 in individual HAL36 (hg H23). a) classical 'Sanger' or dye-terminator sequencing ; b) trace view of probe intensity as consecutive bases from the Affymetrix v2.0. CEL files c) alignment of four different QSTs (1, 3, 6 and 12) against the rCRS (AC_000021).

Fourth, the 25-mer based Mitochip v2.0 resequencing array is based on the rCRS[5]. Due to historical developments, current haplogroup nomenclature places the rCRS in hg H2a2a[5,25,26]. As Supplementary Figure 4 shows, this means that any human mt genome sequence outside hg H2 must necessarily have a minimum of 6 SNP variants compared to the rCRS: at nps 263, 8860, 15326 ('out of hg H2a2a'); np 750 ('out of hg H2a2'); np 4769 ('out of hg H2a'); and np 1438 ('out of hg H2'). Of the 39 hg H archaeological individuals investigated, 38 lie outside hg H2 (Table 1): with the remaining individual (BZH8) belonging to hg H2a1a3. With the two duplicate archaeological samples included (i.e. 41 Mitochip v2.0 runs), this means in total there should be (40 x 6) + (1 x 4) = 244 SNP variants to test the performance and accuracy of the Mitochip v2.0 when used with probes made from ancient DNA libraries enriched for human mtDNA sequences. Of these expected variants, 242/244 gave the correct expected calls for all levels, QST 1, 3, 6 and 12. Two samples, BZH1a and SALZ 57A, gave N at SNP 4769 for QST 1, 3, 6 and 12, due to missing data (i.e. the region including the SNP could not be read). However, the replicate sample of BZH1a (BZH1b) gave the correct 4769G at all levels, QST 1, 3, 6 and 12, showing that separate samples from the same archaeological individual can produce different quality outcomes due to both differential preservation of DNA and stochastic variation.



Supplementary Figure 4. Snippet of mtDNA phylogeny, illustrating the six diagnostic SNPs (highlighted in red) within hg H that define the rCRS as haplotype H2a2a (modified after http://www.phylotree.org/, Build 13).

Fifth, in addition to these generic 'out of H2a2a' SNPs common across our sample set, there were also those pivotal diagnostic SNPs and hierarchies of SNPs that place archaeological individuals into particular established subgroups of hg H according to the most recent release of PhyloTree, Build 14[26] (Table 1). In total there are 159 SNP variants in this category. Of these 159 SNP variants, 96 correspond to previously characterised subhaplogroup-defining SNPs. Many of these Mitochip-identified SNPs belong to root-to-tip hierarchies of phylogenetic branching points. The observation of continuous root-to-tip chains, with no missing SNPs, is another key test of the accuracy and reliability of our strategy to type prehistoric mt genomes via the Mitochip v2.0.

For example, for archaeological individual ESP15, beginning with the rCRS (hg H2a2a) basis of the Mitochip v2.0, the hierarchical flow is thus:

- nps 263, 8860, 15326 (out of hg H2a2a);
- np 750 (out of hg H2a2)
- np 4769 (out of hg H2a)
- np 1438 (out of hg H2) – arriving at the root of hg H
- np 16362 (into pre-hg H6)
- nps 239, 16482 (into hg H6)
- np 3915 (into hg H6a)
- nps 4727, 9380 (into hg H6a1)
- np 11253 (into hg H6a1a)

All these key root-to-tip SNPs were successfully identified via our Mitochip v2.0 run. Moreover, SNPs defining derived twigs on the H6a1a branch at nps 5460 (hg H6a1a1) and 7325, 9362, 11611 and 16311 (hg H6a1a1a) were not detected. This means we can unambiguously place archaeological individual ESP15 within hg H6a1a at a basal position. It is important to bear in mind that there were nine phylogenetic branching points involved in placing individual ESP15 within hg H6a1a, involving 13 positive SNPs and dozens of exclusionary ones.

The remaining 63 (out of 159) SNPs are private, of which 25 are located in the coding region and 38 in the D-loop (Table 1). Lastly, 29 of the latter correspond to 16519C, which is a mutational 'hotspot' in evolutionary terms[27,28], although it appears to be stable and repeatable in individual terms, based on duplicated samples and sequence data derived independently from the Mitochip v2.0.

Overall then, the reliability of the Mitochip v2.0 data is confirmed by the fact that there are no missing links in the hierarchical phylogenetic flow and that there is only one root-to-tip route for every ancient individual. Over the whole sample set (plus duplicates) there were no absences of predicted SNP variants or a 'missing links' at previously characterised phylogenetic branching points. Every archaeological individual occupies a single position on the phylogenetic tree. Along with the independent confirmation of SNPs described above (and below), these data provide strong evidence that using ancient DNA libraries enriched for human mtDNA sequences as probes, the Mitochip v2.0 can generate robust and reliable mt genome sequences and can play a very useful role in the detailed classification of mt haplogroups from prehistoric human populations.

Sixth, six selected enriched libraries were run individually on the Pacific Biosciences platform as described below and SMRT reads were compared to results generated via the Mitochip v2.0. We deliberately chose samples that either had a high number of private SNPs, such as OSH2, BZH4, and BZH6 (four, four and six private SNPs, respectively) or from samples where we observed a background level of SNPs that correspond to the probe (BZH1, DEB9; see section 1.5 for an explanation). All private SNPs observed consistently in the Mitochip v2.0 data (at all QST levels) were confirmed by deep sequencing using the SRMT platform. The mean percentage of majority allele calls for private SNPs (86.2%) is not significantly different (i.e. lower) from the expected (i.e. previously characterised) SNPs (89.9%) when compared for these three samples ($p=0.109$; Nonparametric Wilcoxon Signed Ranks Test), indicating that the allele calls of private SNPs are as robust and reliable as ones that reflect examples previously characterised in the established nomenclature.

Supplementary Table 3. Direct PCR, followed by Sanger sequencing, results for HVS I, HVS II and selected coding region SNPs.

Individual	hg mt genome	HVS I Haplotype	np	HVS II Haplotype	np	Coding SNPs
HAL36	H23	rCRS	15997-16409			10211
HAL11	H	16093C, 16129A	15997-16409			
HAL32	H26	rCRS	15997-16409			11152
HAL39	H1e	rCRS	15997-16409			
DEB9	H88	CRS	15997-16409			
DEB21	H1j	CRS	15997-16409			
KAR 6A	H1bz	rCRS	16046-16402	263G, (315.1C)	34-397	
KAR 11B	H	rCRS	16046-16402	152C, 263G, (315.1C)	34-397	
KAR 16A	H46b	rCRS	15997-16409	263G, (315.1C)	34-397	2772, 11893
OSH2	H89	rCRS	15997-16409			
OSH3	H1	rCRS	16017-16409			
OSH1	H16	rCRS	16056-16409			
OSH7	H5b	16304C	15997-16409			
SALZ18A	H10i	16093C	16046-16402			
SALZ 21B	H1e7	rCRS	15997-16409	263G, (315.1C)	34-397	
ESP30	H1e1a5	rCRS	15997-16401			
HQU4	H7d5	rCRS?	16288-16409			
SALZ57A	H3	rCRS	15997-16409	152C, 263G, (315.1C)	34-397	
SALZ77A	H3	rCRS (ACAD) 16150T (Maina)	15997-16409			
ESP15	H6a1a	16362C	15997-16401			
BZH6	H1ca1	16189C	16056-16193, 16210-16409			3010
BZH4	H1e8	16293G	16056-16409			3010
ROT6	H5a3	16304C	15997-16409			
ALB1	H3b	rCRS	15997-16409			
ROT1	H3ao2	16256T	15997-16409			
ROT2	H5a3	16304C	15997-16409			
QUEXII1	H4a1	rCRS	15997-16409			
QUEXII2	H4a1	rCRS	15997-16409			
QLB 26A	H1	rCRS	16046-16402	263G (309.1C, 309.2C, 315.1C)	34-397	
QUEXII3	H13a1a2c	rCRS	16056-16409			
QLB 28b	H1	rCRS	16046-16402	263G, (309.1C, 309.2C, 315.1C)	34-397	
BZH1	H11a	16293G, 16311C	16056-16409			
BZH8	H2a1a3	16240t, 16354T	16056-16409			
BZH14	H82a	16220G	16056-16409			
EUL 41A	H4a1a1a2	rCRS	15997-16409	73G, 263G, (309.1C, 315.1C)	34-397	
EUL 57B	H3	rCRS	15997-16409	152C, 263G, (315.1C)	34-397	
QUEVIII4	H7h	16213A	15997-16409			
Sardinia	H1aw1	rCRS	15997-16409			
Iron Age	H90a	rCRS	15997-16409			

2.4. Affymetrix Mitochip v2.0 data analysis

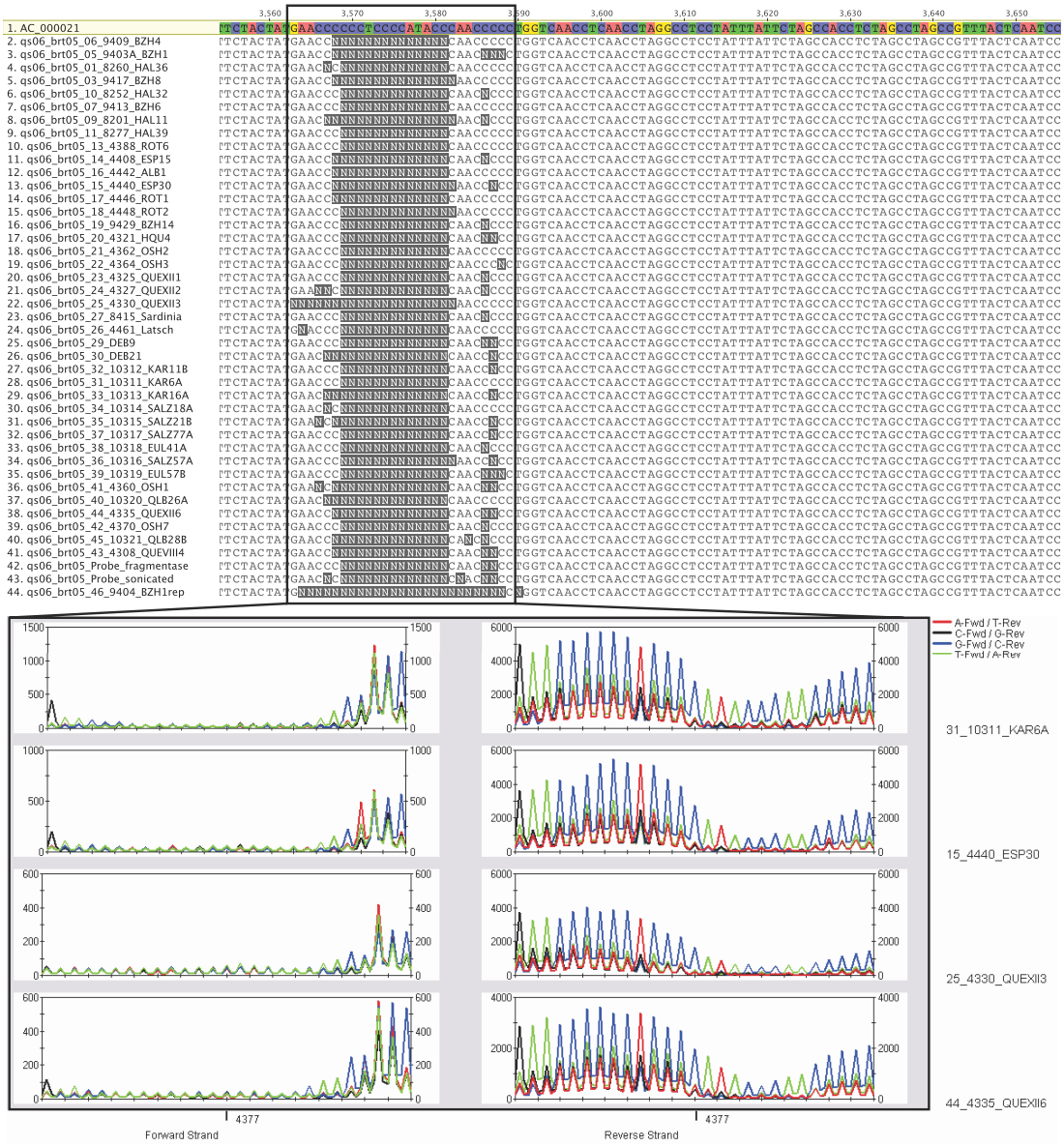
We visually inspected all 16544 nucleotide positions of the GSEQ outputs at QSTs 1, 3, 6 and 12 as an alignment against the rCRS and in parallel examined trace views of probe intensities in the respective trace CEL files. Redundant probes tiled to cover most common control region variants were not analysed.

We identified and characterised regions, which were not, or only poorly, covered (Supplementary Table 4). These can largely be described in two categories, AT-rich regions and poly-cytosine stretches (C-stretches). It is well known that the hybridisation of AT-rich regions of DNA is less efficient due to the lower melting temperature and potential secondary DNA structure of AT-rich regions[29]. Over three consecutive rounds of hybridisation and enrichment, as well as during the final hybridisation to the Mitochip v2.0, certain restricted AT-rich regions were clearly selected against. In addition, even without prior enrichment/hybridisation rounds Hartmann et al.[16] observed that the vast majority of non-called nucleotides (N calls) occurred in regions with a run of ≥4Cs, within a 25-base window of a mutation, and within a 25-base window of ≥15 As and Ts. By visual inspection of all samples using the probe intensity view, we could observe that for all C-stretch regions at least one direction of the tiling arrays (forward or reverse) was readable. We therefore used QST 6 outputs to manually reclaim C-stretch regions, and could decrease the number of N calls to an average of 248, resulting in a final call rate of 98.5% across all samples. This corresponds to a mean coverage of 98.5% of the mt genome considering the 12 bases at the beginning and end of the mt genome that are not interrogated by the Mitochip v2.0 (Supplementary Table 5). Judging from the probe intensity view in GSEQ, AT-rich regions are usually not well covered in both directions, indicating that these regions are successively lost/decreased during hybridisation rounds (including the Mitochip), whereas C-stretches are well covered, but can generate readability difficulties on one strand within the 25 base window that is interrogated by the Mitochip v2.0 tiling array (Supplementary Figure 5).

We realised that by individual visual inspection of probe intensities and by applying signal intensity thresholds, the number of remaining N or discordant calls could be further reduced, as had been described recently[30]. In our case, however, under the assumption that the final round enriched library does not only contain 100% human mtDNA sequences (and damage derived base changes in a sub-fraction of these), but also a proportion of potentially similar microbial analogous sequences, we chose a conservative approach and kept QST 6 outputs, that were manually corrected for C-stretch region only, as final calls.

The final designation of mitochondrial genome sequences into sub-haplogroups of hg H – via identified SNP differences to the rCRS – was therefore the result of several lines of

evidence: the Mitochip v2.0 runs themselves (with a comparison between the QST outputs at 1, 3, 6 and 12); independent direct PCR and Sanger sequencing of previously characterised and novel 'private' SNPs from DNA extracts and via SMRT sequencing; and a final phylogenetic validation using the – at present – largest available dataset for hg H mt genomes (provided by DMB and LQ). The rationale of the PhyloTree database is to list subgroup-defining SNPs if they have been observed at least twice from two independent, unrelated sequences/individuals. Most of the end-tip sequences on the mt phylogeny contain a number of additional private SNPs that are not listed in the tree, but are potentially informative for a higher resolution pending future confirmation as additional sequences become available. In the light of the ever-growing nature of this database, the status of these private SNPs will remain tentative for now. However, the fact that all 16 private SNPs that were observed on the Mitochip v2.0 could unambiguously be confirmed by next generation sequencing, allows us to conclude that the detection of private SNPs can be relied upon when sufficiently high QSTs are applied to MitoChip-derived data. Final mt genome sequences data were deposited in Genbank under acc. XXX-XXX. CEL files are available upon request and will be made available at http://www.adelaide.edu.au/acad/publications/data.



Supplementary Figure 5. Reclaiming cytosine-stretches for final mt genome sequence calls. Alignment of QST 6 output for all individuals illustrates N calls for C-stretches, whereas in all cases the tiling array for the reverse strand reveals readable data (zoom in for four selected samples).

Supplementary Table 4. List of problematic regions of the human mt genome as defined by N calls from an alignment of all archaeological mt genome sequences at QST 6.

Problematic regions by np	location	description
208	D-loop	AT-rich
301-312	D-loop	C-stretch
435-436	D-loop	C-stretch
495-501	D-loop	C-stretch
803-805	RNR1	C-stretch
956-965	RNR1	C-stretch
1683-1684	RNR2	C-stretch
2490	RNR2	C-stretch
3169-3171	RNR2	C-stretch
3211	RNR2	C-stretch
3485-3486	ND1	C-stretch
3527-3529	ND1	C-stretch
3568-3586	ND1	C-stretch
3894-3896	ND1	C-stretch
3960-3962	ND1	C-stretch
4138-4140	ND1	C-stretch
4251-4252	ND1	C-stretch
4761-4775	ND2	AT-rich
5084-5094	ND2	AT-rich
5233-5235	ND2	C-stretch
5304-5305	ND2	C-stretch
5439	ND2	C-stretch
5450	ND2	C-stretch
5495-5517	ND2/TRNW	AT-rich
5897	non-coding	C-stretch
6314	COX1	C-stretch
6568	COX1	C-stretch
6847	COX1	C-stretch
7297-7312	COX1	AT-rich
7332-7334	COX1	unclear
7399-7406	COX1	C-stretch
7468-7469	TRNS	C-stretch
7492-7493	TRNS	C-stretch
7517-7535	non-coding/TRND	AT-rich
7540-7569	TRND	AT-rich
7817	COX2	C-stretch
7819	COX2	C-stretch
8030	COX2	C-stretch
8496-8505	ATP8	AT-rich
8560-8561	ATP8/ATP6	C-stretch
9527-9528	COX3	C-stretch
9556-9558	COX3	C-stretch
9573	COX3	C-stretch
9909	COX3	unclear
9998-10005	COX3	AT-rich
10112-10113	COX3	AT-rich
10193-10195	ND3	C-stretch
10280	ND3	C-stretch
10450-10459	TRNR	AT-rich

Problematic regions by np	location	description
10487-10495	ND4L	AT-rich
10675-10676	ND4L	unclear
10939	ND4	C-stretch
10949-10950	ND4	C-stretch
11142	ND4	C-stretch
11235-11236	ND4	C-stretch
11377-11380	ND4	AT-rich, palindrome
11428-11249	ND4	C-stretch
11675	ND4	C-stretch
11869	ND4	C-stretch
12086-12087	ND4	C-stretch
12110	ND4	C-stretch
12387	ND5	C-stretch
12970	ND5	C-stretch
13027-13028	ND5	C-stretch
13054-13060	ND5	C-stretch
13648-13650	ND5	C-stretch
13678-13679	ND5	C-stretch
13755-13765	ND5	C-stretch
14063	ND5	CA-rich
14156-14158	ND5	C-stretch
14245-14247	ND5	C-stretch
14342	ND6	C-stretch
14419-14427	ND6	C-stretch
14491-14515	ND6	C-stretch/AT-rich
14532	ND6	C-stretch
14777-14781	CYTB	AT-rich
14809-14816	CYTB	C-stretch
15265-15268	CYTB	C-stretch
15527-15528	CYTB	C-stretch
15538-15544	CYTB	C-stretch
16186-16191	D-loop	C-stretch
16260-16262	D-loop	C-stretch
16377	D-loop	C-stretch

Supplementary Table 5. Mitochip v2.0 call rates (in percentages) and absolute numbers of bases not called per ancient mt genome under different quality threshold settings.

ACAD	Individual	Callrate QST1	Ns QST1	Callrate QST3	Ns QST3	Callrate QST6	Ns QST6	Callrate QST12	Ns QST12	Ns after editing QST6	Final
8260	HAL36	99.5	90	98.7	220	97.9	351	96.7	548	172	99.0
9417	BZH8	99.3	118	98.4	259	97.6	396	96.5	578	180	98.9
9403	*BZH1*	98.7	*215*	97.9	*355*	96.8	*525*	95.3	*780*	318	98.1
9409	BZH4	99.3	121	98.5	247	97.6	404	96.0	658	214	98.7
9413	BZH6	99.2	131	98.3	278	97.6	403	96.3	611	201	98.8
8201	HAL11	99.3	114	98.3	277	97.4	425	96.2	632	205	98.8
8252	HAL32	99.3	117	98.6	227	97.8	362	96.6	563	181	98.9
8277	HAL39	99.4	100	98.7	222	97.9	350	96.7	549	174	98.9
4388	ROT6	99.3	110	98.6	232	97.8	371	96.5	571	179	98.9
4408	ESP15	99.2	125	98.4	270	97.5	408	96.4	601	224	98.6
4440	ESP30	99.2	135	98.4	272	97.4	426	96.3	618	251	98.5
4442	ALB1	99.3	121	98.5	240	97.8	365	96.7	552	199	98.8
4446	ROT1	99.2	132	98.4	263	97.6	389	96.6	558	221	98.7
4448	ROT2	99.5	88	98.7	218	97.8	366	96.7	550	186	98.9
9429	BZH14	99.0	168	98.2	305	97.2	465	95.7	718	262	98.4
4321	HQU4	99.0	165	98.1	312	97.2	462	96.0	661	290	98.2
4362B	OSH2	99.0	159	98.4	267	97.6	402	96.5	575	207	98.7
4364A	OSH3	99.0	162	98.1	307	97.3	444	96.2	631	250	98.5
4325A	QUEXII1	98.9	176	98.2	302	97.4	437	96.2	632	243	98.5
4327B	QUEXII2	98.9	174	97.9	341	96.9	509	95.5	750	274	98.3
4330C	QUEXII3	98.8	204	97.8	368	96.9	518	95.2	797	273	98.3
4461A	Iron Age	99.2	133	98.4	268	97.5	420	96.2	624	195	98.8
8415A	Sardinia	99.3	113	98.5	240	97.8	369	96.7	545	213	98.7
Mainz	DEB9	99.2	125	98.4	271	97.5	412	96.4	598	240	98.5
Mainz	DEB21	99.1	141	98.4	258	97.6	404	96.4	594	228	98.6
10311	KAR 6A	99.5	86	98.7	209	97.9	340	96.8	523	177	98.9
10312	KAR 11B	99.3	113	98.7	215	98.0	338	96.8	524	172	99.0
10313	KAR 16A	98.8	203	97.8	357	97.0	500	95.7	717	303	98.2

ACAD	Individual	Callrate QST1	Ns QST1	Callrate QST3	Ns QST3	Callrate QST6	Ns QST6	Callrate QST12	Ns QST12	Ns after editing QST6	Final
10314	SALZ 18A	99.4	102	98.6	235	97.8	359	96.8	528	183	98.9
10315	SALZ 21B	98.7	210	97.9	352	97.0	494	95.7	707	308	98.1
10316	SALZ 57A	98.3	286	97.1	475	96.2	635	94.4	923	449	97.3
10317	SALZ 77A	99.0	165	98.2	297	97.4	425	96.3	616	246	98.5
10318	EUL 41A	99.1	146	98.3	273	97.6	397	96.4	593	228	98.6
10319	EUL 57B	99.2	140	98.4	269	97.5	413	96.3	618	216	98.7
10320	QLB 26A	99.3	115	98.6	228	97.8	356	96.7	541	191	98.8
4360B	OSH1	98.9	174	98.0	324	97.2	465	95.9	674	276	98.3
4370B	OSH7	99.4	98	98.7	221	98.0	338	96.8	536	176	98.9
4308B	QUEVIII4	98.9	176	98.2	299	97.4	436	96.1	639	270	98.4
4335B	QUEXII6	99.3	112	98.6	225	97.9	353	96.8	532	190	98.9
10321	QLB 28b	98.9	185	98.1	315	97.2	466	95.9	674	289	98.3
9404A	BZH1	96.6	565	94.7	883	92.9	1171	90.4	1590	799	95.2
PROBE	fragmentase	99.3	115	98.4	265	97.5	415	96.3	607	201	98.8
PROBE	sonicated	98.8	195	97.8	371	96.7	549	94.9	848	251	98.5
Average		**99.1**	**154**	**98.2**	**294**	**97.4**	**441**	**96.1**	**650**	**248**	**98.5**

3. Pacific Biosciences SMRT® sequencing and data analyses

Six of the mt-enriched DNA libraries were converted to SMRTbell template libraries for sequencing on a Pacific Biosciences *RS*. First, the samples were blunted by incubating 1µg of enriched amplicons with 1X Template Prep Buffer (Pacific Biosciences, Menlo Park, CA), 1mM ATP, 0.4mM dNTPs, 0.2mg/mL rabbit serum albumin, 20U T4 Polynucleotide Kinase (NEB, Ipswich, MA), 4.5U T4 DNA Polymerase (NEB) and water to 40µL for 15 minutes at 25°C. The blunted samples were then purified with GenCatch spin columns (Epoch Life Science) and A-tailed. To A-tail, samples were incubated in solutions containing the blunted amplicons, 1x Template Prep Buffer, 0.4mM dATP, 8U Klenow Fragment exo– (NEB) and water to 40µL for 60min at 37°C. After tailing, the samples were GenCatch purified and circularized by ligation to hairpin adaptors. Ligations were carried out in reactions containing the A-tailed amplicons, 1X Fermentas T4 Ligase Buffer, 60U T4 Ligase (Fermentas), 0.5µM T-tailed hairpin adaptor (Pacific Biosciences), and water to 60µL. All ligations were incubated at 25°C for 16 h and then heated at 65°C for 10min to inactivate the ligase. Each sample was then treated with 20U of Exonuclease I (NEB) and 10U Exonuclease III (NEB) for 60min at 37°C to digest linear, non-circularised, DNA molecules. Finally, the samples were GenCatch purified and an aliquot of each was assayed on a Bioanalyzer using a High Sensitivity DNA chip.

Single-Molecule, Real-Time (SMRT) sequencing was carried out on the PacBio *RS* at the Pacific Biosciences lab (Menlo Park, USA). Reads were processed and mapped to the respective reference sequences for each plasmid using the Basic Local Alignment with Successive Refinement (BLASR) mapper (http://www.pacbiodevnet.com/smrtanalysis/software/blasr) and the Pacific Biosciences SMRT Analysis pipeline using the standard mapping protocol (http://www.pacbiodevnet.com/smrtanalysis/software/smrtpipe). Circular consensus reads[31] were trimmed to remove the library PCR adapter sequences before mapping to the rCRS (NC_012920) using a maxScore = -50 filter and exported as SAM.files (Supplementary Table 6).

At ACAD, SAM.files were imported into Geneious Pro 5.5.6[32] and re-assembled to the rCRS discarding reads <20bp in length and using a conservative mapping method. Custom sensitivity settings were: 10% maximum gaps per read and a maximum gap size of 3, 20 bases minimum overlap with a word length of 14 (ignoring repeats >5 times) and index word length of 12, 10% maximum mismatches per read and a maximum ambiguity of 4 bases (Supplementary Table 6).

To estimate the ratio of endogenous human mtDNA, we calculated the average rate of redundancy at expected SNP sites, i.e. we counted the number of reads that corresponds to the allele (or the SNP variant) present in the Mitochip data from the same individual (Supplementary Data 1). Average results across all six samples indicate that the majority of the reads (83.8%) were consistently derived from a single endogenous source. This number increases when two samples containing traces of probe DNA sequence are removed (87.4%). The second largest fraction of the reads (11%, and 10.6% for four samples) corresponds to any other state (transitions, transversions and indels) at these particular nps and incorporates varying levels of: a) potential background contamination of exogenous human mtDNA; b) post-mortem miscoding lesion DNA damage; and c) read inaccuracies. Finally, we also calculated the proportion of potential contamination with the probe by checking the 29 SNPs that distinguish our probe (haplotype J1c8) from the rCRS. Of these, six SNPs leading out of the H2 branch are shared with all archaeological individuals (except individual BZH8), and SNP 3010A is shared mutation between subhaplogroups J1 and H1 and are therefore not informative, leaving 22 nps that could be checked for the presence of reads that share an identical allele with the probe. The average proportion of reads that resemble the probe is 1.9% across all four samples, for which we did not have an indication from the Mitochip v2.0 data, and higher (5.1%) when BZH1 and DEB9 were included. This proportion can be considered a maximum, since the coverage is not evenly spread across the mt genome and appears random, and secondly, DNA damage, esp. C>T and G>A transition could mimic read inaccuracies. In addition, the potential contribution of the probe is significantly lower than the average background ($p=0.028$, Nonparametric Wilcoxon Signed Ranks test), therefore not affecting the overall allele call and the final haplogroup designation. With updated probe preparation and mt-enriched library recovery protocols (see above), probe background can be reduced to undetectable levels.

SAM and Assembly files are available upon request and will be made available at http://www.adelaide.edu.au/acad/publications/data.

Supplementary Table 6. Summary of SMRT sequencing data and read assembly.

	BLASR assembly							Geneious Pro re-assembly (20bp cut-off)			
Sample	**BLASR no.**	**BLASR**	**BLASR**	**Unused**	**Unused**	**Unused**	**BLASR**	**Reassembly**	**Reassembly**	**Coverage**	**Mean**
H1_BZH1	51,945	39	68	16,410	43	66	95.9	17,543	34,402	98.5	110 (0/724)
H2_BZH4	44,040	41	66	6,551	41	63	96.9	10,072	33,968	99.1	99 (0/449)
H3_BZH6	62,646	43	69	8,672	41	62	96.8	12,839	49,807	99.8	154 (0/632)
H4_OSH2	46,802	41	67	4,791	38	63	96.2	7,883	38,919	98.7	110 (0/548)
H5_QUEXII3	45,093	43	75	9,039	45	73	95.5	12,930	32,163	99	107 (0/1500)
H6_DEB9	51,808	42	68	6,324	37	63	96.7	8,720	43,088	99.5	125 (0/632)

4. Population genetic analyses

4.1. Network analyses

A median joining network of all ancient hg H samples (Figure 1a) was generated/constructed manually using the most up-to-date version of the mitochondrial phylogenetic tree (PhyloTree.org, mtDNA tree Build 14) as a scaffold on which to place the observed hg H lineages[26,33]. The latest version includes a revised version of the hg H subtree comprising 1203 sequences in total. As per convention, insertions at np 309.1C(C), 315.1C, 523-524d (aka 522-523d), 16182C, 16183C, 16193.1C(C) and mutation 16519 were not considered for phylogenetic reconstruction[26]. Prior to the network reconstruction, the ancient hg H samples were pooled into different temporal/cultural groups in order to calculate genetic diversity indices and to test for genetic differentiation.

4.2. Principal component analysis

Principal component analysis (PCA) is a commonly used form of data reduction, suitable for extracting information from genetic markers (here sub-haplogroups of human mtDNA hg H), as the analysis attempts to display all variation within the data on a small number of axes, while limiting a priori assumptions about the data being analysed. PCA was used to describe and visualise the maternal genetic relationship amongst the Neolithic cultures investigated and to 37 modern-day European and Near Eastern populations (Figure 2a). PCA was performed on the frequency of sub-clades of mtDNA hg H taken from the literature and is summarised in Supplementary Table 7. To minimize statistical noise caused by rare sub-clades and to allow for data compatibility across published studies we considered only the following 15 most common hg H sub-clades in Europe and the Near East: H*, H1, H1a, H1b, H2, H2a1, H3, H4, H5, H5a, H6, H6a, H7, H8 and H11. PCAs were performed and visualized in R version 2.11.1[34] using a customized script based on the function *prcomp* freely available upon request.

Ancient hg H samples were pooled into three different groups based on number of samples available for ‘pan-European’ archaeological phenomena/cultures alongside hypothesized geographic origins (LBK, n=9 and BBC, n=7), and a temporally transitional group pooling regional (mostly Mid-Neolithic) cultures (MNE, n=10) due to small sample size, excluding distinct but small sample sets such as the Corded Ware (n=2) and later Bronze Age Unetice (n=5) data. This MNE pool clustered with present-day Central European populations (Figure 2), signifying the presence of a substratum of hg H diversity in the Mid-

Neolithic different from certain prominent cultures such as the preceding LBK and the subsequent Bell beakers. This might be explained by a phase of more localised/regional development after the initial Neolithisation, even allowing for moderate influences from contemporaneous Mid-Neolithic cultures outside our region[35,36]. We note however that the observed clustering may have reflected an artificial homogenisation caused by grouping many potentially distinct supra-regional influences from diverse Mid-Neolithic cultures.

Supplementary Table 7. Details of comparative Hg H sub-clade frequency data used for principal component analyses.

Population	abbreviation	n	H*[1]	H1	H1a	H1b	H2a	H2a1	H3	H4	H5	H5a	H6	H6a	H8	H7	H11
Galicia[37]	GAL	124	15.3[2]	35.5	2.4	0.0	4.0	2.4	20.2	7.3	0.0	6.5	0.0	4.8	0.8	0.8	0.0
Cantabria[37]	CNT	53	11.3	37.7	7.5	0.0	5.7	0.0	17.0	5.7	3.8	5.7	0.0	5.7	0.0	0.0	0.0
Catalonia[37]	CAT	40	22.5	32.5	2.5	2.5	0.0	2.5	12.5	5.0	5.0	10.0	0.0	5.0	0.0	0.0	0.0
Galicia/Asturia[38]	GAS	65	23.1	41.5	0.0	1.5	9.2	0.0	10.8	1.5	1.5	6.2	1.5	3.1	0.0	0.0	0.0
Cantabria2[38]	CAN	31	16.1	41.9	0.0	3.2	0.0	0.0	16.1	0.0	6.5	3.2	0.0	12.9	0.0	0.0	0.0
Cantabria3 (Potes)[38]	POT	38	21.1	57.9	0.0	2.6	0.0	0.0	18.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Cantabria4 (Pasiegos)[38]	PAS	22	9.1	63.6	9.1	0.0	0.0	0.0	0.0	18.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Basques (Vizcaya)[38]	VIZ	37	8.1	59.5	0.0	5.4	2.7	0.0	10.8	8.1	0.0	2.7	2.7	0.0	0.0	0.0	0.0
Basques (Guipuzcoa)[38]	GUI	63	7.9	42.9	0.0	3.2	20.6	1.6	19	0.0	0.0	4.8	0.0	0.0	0.0	0.0	0.0
Basques[38]	BMI	85	14.1	43.5	0.0	3.5	9.4	1.2	17.6	2.4	1.2	7.1	0.0	0.0	0.0	0.0	0.0
Iberian Peninsula NE[38]	IPNE	52	26.9	28.8	0.0	7.7	9.6	1.9	7.7	1.9	0.0	3.8	0.0	3.8	0.0	5.8	1.9
Turkey[39]	TUR	90	58.9	10.0	2.2	1.1	2.2	1.1	0.0	5.6	7.8	2.2	1.1	0.0	1.1	5.6	1.1
Armenia[39]	ARM	54	57.4	7.4	0.0	0.0	1.9	7.4	0.0	9.3	3.7	1.9	3.7	0.0	1.9	5.6	0.0
Georgia[39]	GEO	30	60.0	6.7	0.0	0.0	0.0	3.3	0.0	3.3	23.3	0.0	0.0	0.0	0.0	3.3	0.0
Northwest Caucasus[39]	NWC	69	59.4	13.0	0.0	1.4	1.4	4.3	1.4	0.0	8.7	0.0	1.4	4.3	0.0	2.9	1.4
Dagestan[39]	DAG	69	39.1	10.1	0.0	0.0	29.0	5.8	1.4	2.9	0.0	1.4	2.9	4.3	1.4	1.4	0.0
Ossetia[39]	OSS	45	62.2	11.1	0.0	2.2	0.0	2.2	0.0	6.7	8.9	0.0	0.0	4.4	0.0	0.0	2.2
Syria[39]	SYR	28	57.1	0.0	0.0	0.0	10.7	0.0	0.0	3.6	7.1	0.0	7.1	3.6	7.1	3.6	0.0
Lebanon[39]	LBN	34	50.0	20.6	0.0	0.0	2.9	2.9	0.0	2.9	11.8	5.9	0.0	0.0	0.0	0.0	2.9
Jordan[39]	JOR	33	75.8	9.1	0.0	0.0	3.0	3.0	0.0	3.0	6.1	0.0	0.0	0.0	0.0	0.0	0.0
Arabian Peninsula[39]	ARB	50	48.0	4.0	0.0	0.0	2.0	14.0	0.0	10.0	4.0	0.0	12.0	4.0	0.0	2.0	0.0
Arabian Peninsula2[40]	ARE	24	75.0	4.2	0.0	0.0	0.0	0.0	0.0	0.0	4.2	0.0	4.2	4.2	0.0	0.0	8.3
Karachay-Balkaria[39]	KBK	50	38.0	0.0	8.0	10.0	4.0	8.0	4.0	0.0	24.0	0.0	0.0	0.0	0.0	2.0	2.0
Macedonia[40]	MKD	82	50.0	12.2	0.0	3.7	0.0	2.4	1.2	6.1	12.2	2.4	0.0	3.7	0.0	4.9	1.2
Volga-Ural region[41]	VUR	50	42.0	28.0	6.0	0.0	0.0	0.0	0.0	4.0	0.0	0.0	4.0	0.0	0.0	8.0	8.0
Finland[41]	FIN	31	48.4	6.5	12.9	0.0	6.5	6.5	6.5	0.0	3.2	9.7	0.0	0.0	0.0	0.0	0.0
Estonia[41]	EST	50	38.0	20.0	4.0	12.0	2.0	6.0	6.0	0.0	2.0	2.0	4.0	0.0	0.0	2.0	2.0
Eastern Slavs[41]	ESV	165	31.5	21.8	4.2	5.5	3.0	8.5	4.2	1.8	1.8	4.2	6.1	0.0	0.0	1.2	6.1

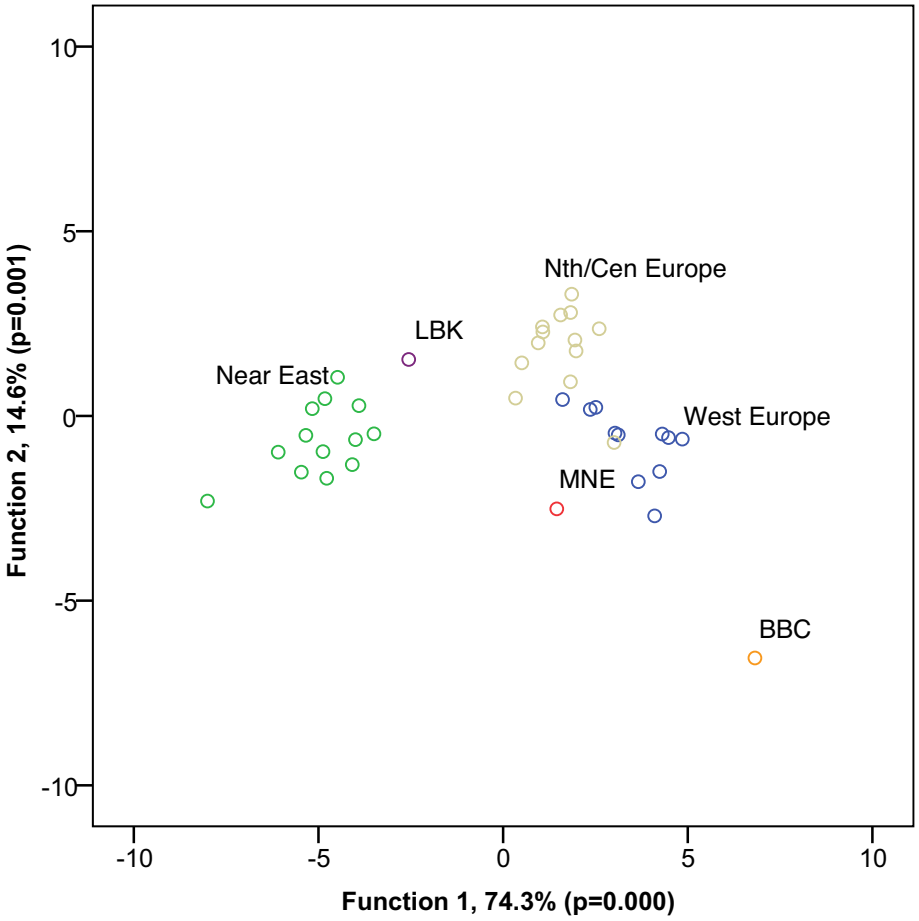
Population	abbreviation	n	H*[1]	H1	H1a	H1b	H2a	H2a1	H3	H4	H5	H5a	H6	H6a	H8	H7	H11
Slovakia[41]	SVK	50	40.0	6.0	8.0	4.0	0.0	2.0	4.0	4.0	8.0	2.0	6.0	0.0	2.0	2.0	12.0
France[41]	FRA	50	34.0	22.0	4.0	0.0	4.0	0.0	12.0	0.0	8.0	2.0	6.0	0.0	0.0	8.0	0.0
Balkans[41]	BLK	50	46.0	8.0	0.0	4.0	0.0	0.0	8.0	0.0	6.0	10.0	8.0	0.0	0.0	6.0	4.0
Germany[40]	DEU	26	42.3	11.5	3.8	7.7	0.0	0.0	7.7	0.0	7.7	7.7	0.0	7.7	3.8	0.0	0.0
Austria[40]	AUT	959	41.0	20.6	3.3	3.8	4.0	2.4	7.2	2.5	4.3	4.8	0.3	2.9	0.1	2.4	0.4
Romania[40]	ROU	102	21.6	26.5	6.9	4.9	0.0	0.0	0.0	1.0	12.7	4.9	0.0	2.9	0.0	16.7	2.0
France Normandy[38]	FRM	37	37.8	29.7	5.4	2.7	0.0	0.0	10.8	0.0	2.7	0.0	2.7	5.4	2.7	0.0	0.0
Western Isles[38]	WIS	39	20.5	35.9	0.0	5.1	0.0	0.0	12.8	5.1	2.6	5.1	0.0	7.7	0.0	2.6	2.6
Czech Republic[38]	CZE	31	22.6	32.3	3.2	3.2	0.0	0.0	6.5	0.0	6.5	9.7	0.0	9.7	0.0	0.0	6.5
Linear Pottery culture	LBK	9	66.7	33.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bell Beaker	BBC	7	14.3	28.6	0.0	0.0	0.0	0.0	28.6	14.3	0.0	14.3	0.0	0.0	0.0	0.0	0.0
Middle Neolithic	MNE	10	30.0	30.0	0.0	0.0	0.0	0.0	20.0	0.0	10.0	0.0	0.0	0.0	0.0	10.0	0.0

[1]binning category summarizing all remaining H sub-clades

[2]values given as percentage %

4.2.1. Discriminant function analysis

To test whether the 'geographic' spread of the PCA plot is real or a random artifact of frequency data, we performed a linear functional discriminant analysis in SPSS version 17.0. This analysis was performed using the PCA scores from the first 15 principal components. The geographic spread observed in the PCA plot was reproduced by the functional discriminant analysis, as seen in the plot of function 1 and 2 (Supplementary Figure 6). Furthermore, the functional discriminant analysis significantly supported the grouping of the Early (LBK) and Middle/Late Neolithic (BBC) cultures with the modern Near Eastern and North/Central and Western European populations, respectively.



Supplementary Figure 6. Linear functional discriminant analysis of PCA scores.

4.2.2. Multivariate analysis of variance (MANOVA)

To confirm whether the clustering pattern observed in the PCA between the ancient and modern populations was significantly supported, we performed a non-parametric multivariate analysis of variance (NP-MANOVA). The NP-MANOVA was performed on a Raup-Crick distance matrix, which was produced from the presence/absence of the 15 hg H sub-clades used in the PCA. Calculations were performed in PAST version 2.09 with 10,000 permutations per test and post-hoc Bonferroni correction to account for multiple comparisons[42]. In accordance with the linear functional discriminant analysis, the NP-MANOVA results significantly supported the presence of three distinct geographic groups from the genetic data (Supplementary Tables 8a). We also tested whether the temporal grouping of ancient individuals in archaeological time periods, which ranged from a narrow to a broader definition, would be statistically supported. Here, only the latter, i.e. lumping both the Mid-Neolithic samples and the Bronze Age samples into broader 'Late Neolithic' pool was significantly supported by the NP-MANOVA (Supplementary Tables 8b-d).

Supplementary Table 8a-d. Summary of NP-MANOVA tests of a) cultures grouped with geographic regions, and between Neolithic samples pooled in different time periods: b) four time periods; c) three times periods; and d) two time periods. Bonferroni corrected values are given in bold print and areas shaded grey indicate significant values ($p<0.05$).

a) Cultures grouped with geographic regions as in Figure 2.

	Iberia	Near East	Mainland Europe
Iberia	0	0	0.0001
Near East	**0**	0	0.0004
Mainland Europe	**0.0003**	**0.0012**	0

b) Four time periods

	Early Neolithic	Middle Neolithic	Late Neolithic	Bronze Age
Early Neolithic	0	0.1456	0.0497	0.0355
Middle Neolithic	**0.7734**	0	0.9168	0.7449
Late Neolithic	**0.2652**	**1**	0	0.5289
Bronze Age	**0.234**	**1**	**1**	0

c) LBK, BBC, and pooled intermediate Neolithic (MNE) as used in PCA

	LBK	MNE	BBC
LBK	0	0.2463	0.0912
MNE	**0.734**	0	0.7961
BBC	**0.2901**	**1**	0

d) Two time periods

	Early Neolithic	Late Neolithic
Early Neolithic	0	0.0034
Late Neolithic	**0.0034**	0

4.3. Population genetic summary statistics and mismatch distributions

All population genetic statistics were calculated in Arlequin version 3.5[43]. We used jMODELTEST 0.1.1[44] in order to find the best fitting evolutionary model and, if required, to estimate a discrete gamma shape parameter for our 39 non-partitioned mt genomes. Based on the resulting scores for each model (AIC and BIC), we subsequently used the Tamura & Nei model and a gamma value 0.049 for our calculations of population distances in Arlequin.

Pairwise population comparison via F_{ST} values showed that genetic distances increased with time during the duration of the Neolithic, reaching the highest and significant values (p=0.00386) between the Early Neolithic and the earliest Bronze Age (Supplementary Table 9a).

Supplementary Table 9a-c. Population pairwise and linearised Slatkin's F_{ST}s between Neolithic samples pooled in different time periods: a) four time periods, b) two pan-European cultures (LBK, BBC) and one intermediate period (MNE), and c) two time periods. Slatkin F_{ST}s are italicised (upper diagonal) and significant distances are given in bold print.

a) Four time periods

	Early Neolithic	Middle Neolithic	Late Neolithic	Bronze Age
Early Neolithic	0	*0*	*0.03179*	*0.09555*
Middle Neolithic	0.01135	0	*0*	*0.02299*
Late Neolithic	0.02247	-0.01165	0	*0.01148*
Bronze Age	**0.08722**	0.03081	-0.02250	0

b) LBK, BBC, and pooled intermediate Neolithic (MNE) as used in PCA

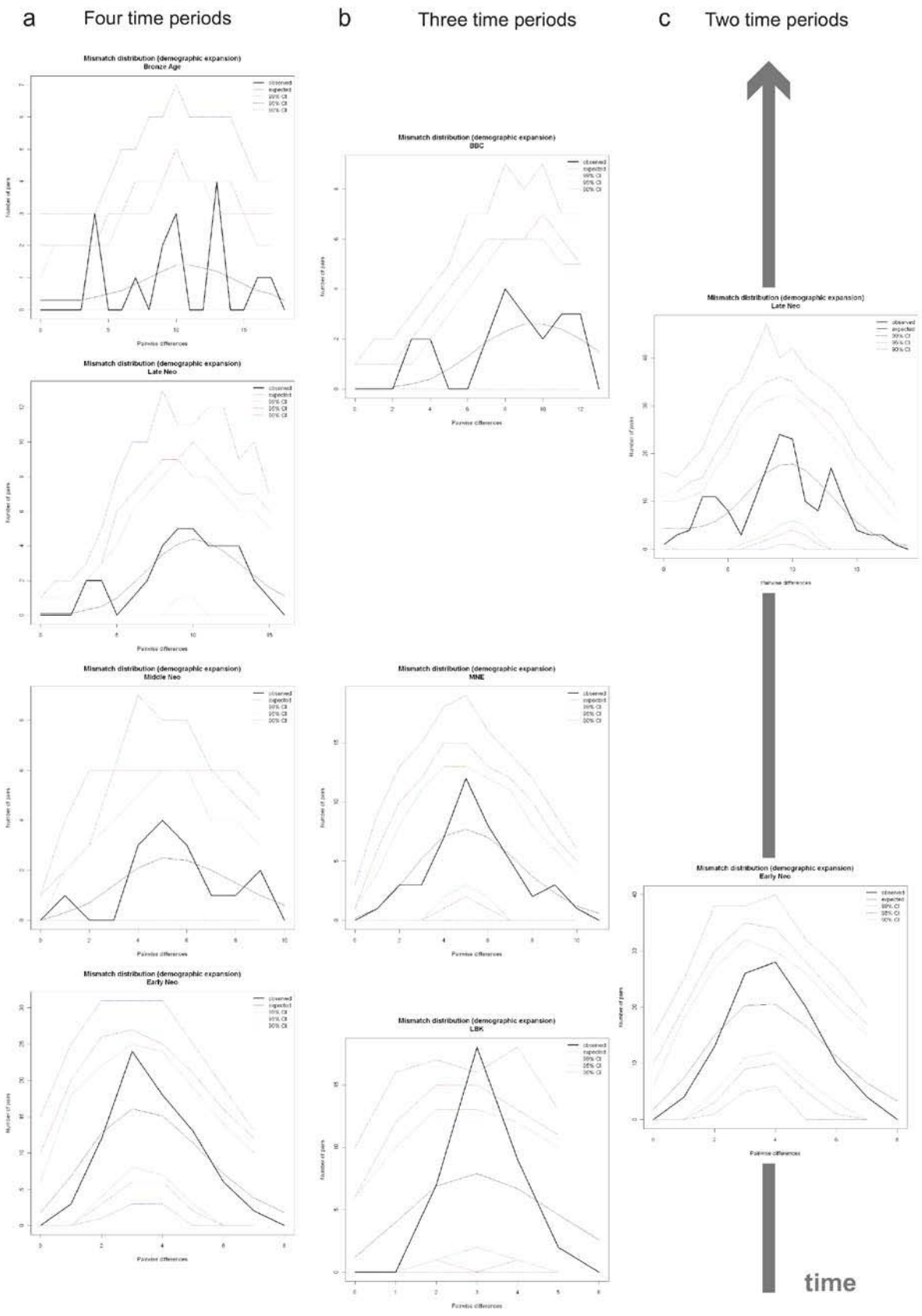
	LBK	MNE	BBC
LBK	0	*0*	*0.03369*
MNE	-0.02587	0	*0*
BBC	0.03260	-0.00704	0

c) Two time periods

	Early Neolithic	Late Neolithic
Early Neolithic	0	*0.01459*
Late Neolithic	0.01438	0

We further calculated mismatch distributions for each of the temporal/cultural groups in order to assess signals of demographic expansion during the Neolithic (Supplementary Figure 7a-c). A uni-modal distribution should indicate a period during or shortly after a demographic expansion, whereas a more 'ragged', multi-modal distribution is commonly found in more stationary populations. When put in temporal order, we could observe a shift from a smooth uni-modal distribution in the early Neolithic and especially in the LBK towards a very ragged multi-modal distribution in later stages of the Neolithic as well as the early Bronze Age[45,46]. An identical pattern could be observed for estimates of spatial expansion based on mismatch distributions (data not shown). Uni-modal mismatch distributions for early Neolithic samples suggesting a recent expansion are consistent with the spread of farming lifestyles during the Neolithic transition.

The tendency towards more stationary populations in the Late Neolithic and Bronze Age suggest a process of stasis and/or diversification and possibly a higher impact of drift in smaller and more restricted settlement areas, in accordance with the increase in pairwise FSTs during the Neolithic.



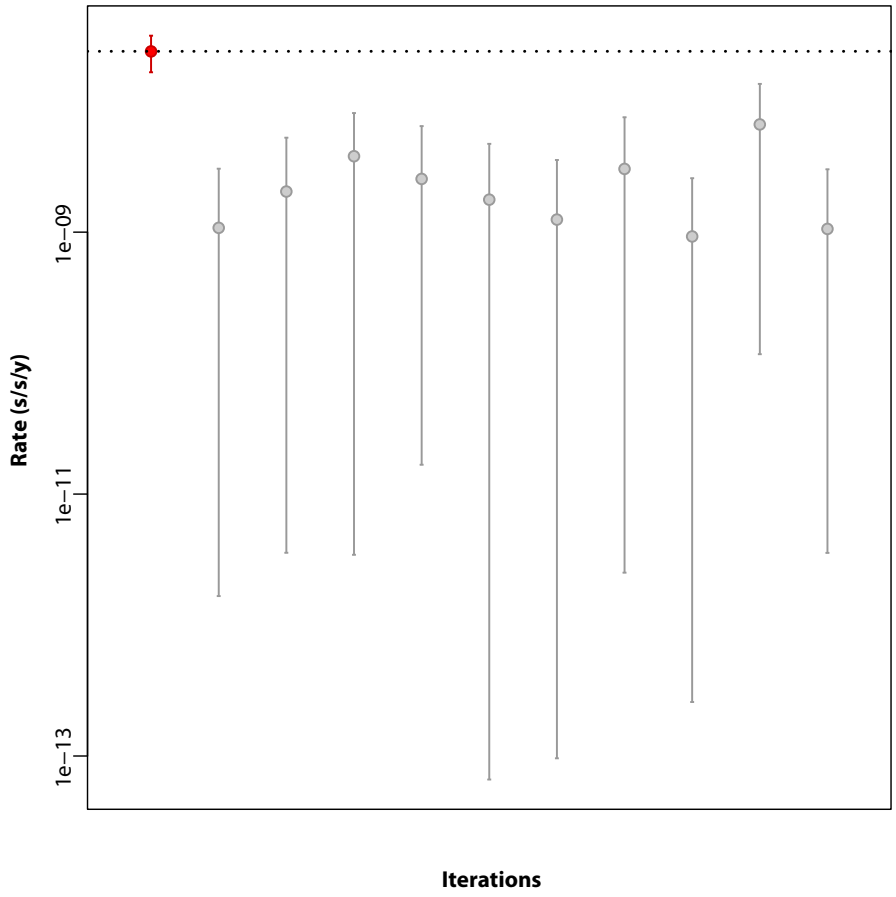
Supplementary Figure 7. Mismatch distribution graphs in temporal order indicate a decrease in demographic expansion during the Neolithic.

4.5. Bayesian skyride analyses and mutation rate

The data set comprised 37 newly sequenced, non-related, ancient mitochondrial genomes and 5 sets of randomly chosen distinct modern-day haplogroup H mitochondrial genomes from Phylotree (http://www.phylotree.org, mtDNA tree Build 12 (20th July 2011)) and 420 newly available hg H sequences[47]. The Sequences were aligned to the revised Cambridge Reference Sequence (rCRS: AC_000021)[5] manually using the program SeaView[48]. The alignment was partitioned into 4 subsets, representing the D-loop, the protein-coding regions (1st+2nd codon positions and 3rd codon position) and a concatenation of tRNA and RNA genes. Insertions at nps 309.1C(C), 315.1C, 523-524d (aka 522-523d), 16182C, 16183C, 16193.1C(C) were not considered for phylogenetic reconstruction and position 16519 was removed from the D-loop subset[26]. The best substitution models were selected using ModelGenerator 0.85[49] by comparison of Bayesian Information Criterion scores: HKY+G for D-loop, TN+G for protein-coding regions and HKY for RNA genes. Considering the short evolutionary timescale being studied (intra-haplogroup), models including a proportion of invariant sites were excluded. A Bayesian skyride analysis[50] was performed using the phylogenetic software BEAST 1.6.1[51] and calibrated using radiocarbon dates from 18 of the ancient individuals and mean archaeological dates for the remaining individuals. This allowed us to achieve a broad temporal coverage for ~3500 years of the Neolithic period in Central Europe (5500–2000 BC) and to generate the most precise demographic reconstruction of hg H. Results were replicated using independent sets of 100 (1x), 200 (3x), and 300 (1x) mitochondrial genomes. A strict molecular clock was used, allowing for a distinct rate in each subset of the alignment. Additional analysis using an uncorrelated log normal relaxed clock to account for potential rate variations could not reject the strict clock assumption. Results from two MCMC chains were combined after 100,000,000 iterations and 10% burn-in. All parameters showed sufficient sampling, indicated by effective sample sizes above 200. Tracer 1.5[52] was used to produce the skyride plot shown in Figure 1b. Our data contradict a recent study based on modern-day mt genomes only, which describes a pronounced population increase at ~7000 BC, interpreted as a Neolithic expansion into Europe, but followed by slow population growth until the present day[53].

We carried out a ‘date randomization test’, to test whether the signal from the radiocarbon dates associated with the ancient sequences is sufficient to calibrate the H haplogroup phylogeny[54]. This test randomises all dates associated with the sequences (including modern ones) and replicates of the phylogenetic analysis as described above. If the structure and spread of the ancient sequences in the tree were sufficient to calibrate the

analysis, the inferred mean rate of the randomised analysis should be significantly different from the rate calculated using the correct association date/sequence. In other words, the 95% HPD of the randomised analysis should not overlap with the mean rate estimated without randomisation. The comparison of estimated rates from the main analysis and from ten replicates with randomised dates presented in Supplementary Figure 8 confirms the presence of sufficient signal provided by dates from the 37 ancient samples to calibrate the tree.



Supplementary Figure 8. Date randomisation test for the dated phylogeny. Estimated rates were replicated ten times, each with randomised dates (in grey). These are significantly different from the rate estimated from the main analysis (in red), confirming the presence of sufficient signal from the ancient tip dates to calibrate the tree.

4.6. Haplogroup designation according to the "Copernican" reassessment of the human mitochondrial tree

Access to complete mt genome data from modern-day samples from two recent studies[33,47] has provided the basis for the phylogenetic network and demographic reconstructions of mtDNA hg H for this study. Several thousand novel mt genomes have extended the human mtDNA phylogeny and have deepened the resolution of hg H to 87 sub-clades in particular. At the same, and during the final stages of this study, a reassessment of the mtDNA tree from its root was published, that recommended the use of the Reconstructed Sapiens Reference Sequence (RSRS)[33] over the traditional use of the revised Cambridge Reference Sequence (rCRS)[5]. While our study was conducted using the traditional rCRS, we also report our ancient mt genome data using the new RSRS, anticipating a rapid reception of the new reference and reporting formats (Supplementary Table 10).

When compared to the phylogenetic root (RSRS) all our samples display the following sequence variants leading into branch hg H (in numerical order): G73A (back mutation A73G in EUL41A), C146T, C152T (back mutation T152C! in KAR11B, SALZ57A, EUL57B, OSH1), C182T/T182C!, C195T, (back mutation T195C! in BZH14), A247G, 522.1AC, A769G, A825t, A1018G, G2706A, A2758G, C2885T, T3594C, G4104A, T4312C, T7028C, G7146A, T7256C, A7521G (not covered due to AT-rich region, see section 2.4), T8468C, T8655C, G8701A, (back mutation A8701G! in Sardinia), C9540T, G10398A, T10664C, A10688G, C10810T, C10873T, C10915T, A11719G, A11914G, T12705C, G13105A, G13276A, T13506C, T13650C (T13650N in BZH1, HAL32, HAL39, ALB1, HQU4, OSH1, QUEXII6, QUEVIII4 due to C-stretch, section 2.4), T14766C, G15301A/A15301G!, A16129G, (back mutation G16129A! in HAL11), T16187C, (T16187N in HAL11, ROT1, OSH3, DEB9, KAR6A, QUEXII6 due to C-stretch), C16189T (back mutation T16189C! in BZH6), T16223C, G16230A, T16278C and C16311T.

Supplementary Table 10. Summary of genotyping data using the "Copernican" approach. Sub-haplogroup definig SNPs are shown in bold and 'private' / as-yet-unknown sequence variants in regular print.

Individual	Hg[1]	Hg H sequence variants compared to RSRS
HAL36	**H23**	**C10211T**
HAL11	**H**	T16093C, G16129A!
HAL32	**H26**	**T11152C**
HAL39	**H1e**	**G3010A**, **G5460A**
DEB9	**H88**	**A8596G**
DEB21	**H1j**	**G3010A**, **T4733C**
KAR6a	**H1bz**	G1719A, **G3010A**, C14380T
KAR11b	**H**	T152C!
KAR16a	**H46b**	**C2772T**, A11893G
OSH2	**H89**	A6932G, C8068T, T12696C,
OSH3	**H1**	**G3010A**
OSH1	**H16**	**T152C!**, **C10394T**
OSH7	**H5b**	**C456T, G5471A, T16304C**, C16519T
SALZ18a	**H10i**	C13503T, **T14470a**, T**16093C**
SALZ21b	**H1e7**	T1766C, **G3010A**, **G5460A**
ESP30	**H1e1a5**	**G3010A**, **G5460A**, (C5960T), **A8512G**, G8865A, **C14902T**
HQU4	**H7d5**	**A4793G**, **C15409T**, G16388A
SALZ57a	**H3**	T152C!, **T6776C**
SALZ77a	**H3**	**T6776C**
ESP15	**H6a1a**	**T239C**, **G3915A**, **A4727G**, **G9380A**, **T11253C**, **T16362C**, **A16482G**, C16519T
BZH6	**H1_TBD**	**G3010A**, A8149G, A9377G, T9467C, A13671G, T14319C, T**16189C!**
BZH4	**H1e7**	**G3010A**, **G5460A**, A15220G, A15401G, A16293G
ROT6	**H5a3**	**C456T**, **G513A**, **T4336C**, **G15884A**, **T16304C**, C16519T
ALB1	**H3b**	**A2581G**, **T6776C**
ROT1	**H3ao2**	C4577T, **T6776C**, **C16256T**
ROT2	**H5a3**	**C456T**, **G513A**, **T4336C**, **G15884A**, **T16304C**, C16519T
QUEXII1	**H4a1**	**C3992T**, **A4024G**, **T5004C**, **G9123A**, **C14365T**, **A14582G**, C16519T
QUEXII2	**H4a1**	**C3992T**, **A4024G**, **T5004C**, **G9123A, C14365T, A14582G**, C16519T
QLB26a	**H1**	**G3010A**
QUEXII3	**H13a1a2c**	**C2259T**, **A4745G**, G9025A, **A13542G**, **C13680T**, **C14872T**, C16519T
QLB28b	**H1**	**G3010A**
BZH1	**H11a**	**T195C!, T961g, T8448C, (G13759A), A16293G, T16311C!** , C16519T
BZH8	**H2a1a3**	**G951A**, **G1438A**, **G4769A**, **C6173T**, **T13095C**, A16240t, **C16354T**, C16519T
BZH14	**H82a**	T195C!, **A16220G**
EUL41a	**H4a1a1a5**	**A73G!**, **C3992T**, **A4024G**, **T5004C**, **G8269A**, **G9123A**, **A10044G**, C13545T,
EUL57B	**H3**	T152C!, **T6776C**
QUEVIII4	**H7h**	**A4793G**, **G16213A**
-	**H1aw1**	**G3010A**, **A8701G!**, C15912T
-	**H90**	C5435T, T8911C, T10237C, T15109C

[1]Haplogroup H designations based on the http://www.phylotree.org mtDNA tree Build 14 (5 April 2012)[26,33].

5. Archaeological background

The Neolithic in Central Europe is an epoch of cultural and social complexity and dynamism, which is reflected in the succession of archaeologically distinct cultural layers[55]. Many of these are temporally interwoven and the assemblages of artefacts of certain cultures can resemble features of preceding ones. At other times, new cultures appear/emerge without a close resemblance to preceding cultures and carry elements in style and/or artefacts which appear 'exotic' (i.e. have their origins outside Central Europe). This succession of distinct cultural layers provides the basis for our main underlying question: whether cultural changes have been triggered or accompanied by genetic/population changes.

Our area of interest, the Mittelelbe-Saale (MES) region of Saxony-Anhalt, Germany, has long been an attractive settlement area, with a tight network of waterways and close to long-established trade routes. Climatically and geologically, the MES region benefits from its location in the rain-shadow of the nearby Harz Mountains, its deposits of ores and salts, and its fertile black and para-brown soils. This explains why this settlement cluster was highly sought after by many successive groups, whereby a range of demographic models and modes of recurrent streams of colonisation can be assumed.

The Neolithic in Central Europe lasted for roughly 3500 years and is commonly divided in an early, intermediate and later phase, the Early Neolithic, the Middle Neolithic and the Late Neolithic. Our rationale was to type samples from all three phases in a 'transect through time'. Archaeological details for each of the individuals and cultures are given in Supplementary Table 11.

Early Neolithic:

The Early Neolithic (5450–3950 BC) in MES is represented by four archaeological cultures: the *Linienbandkeramik* (linear pottery culture or LBK), the *Stichbandkeramik* (stroke ornamented culture), and the following cultures *Rössen* and *Gatersleben*, named after eponymous sites in MES. A detailed archaeological record traces the spread of agriculture via LBK farmers back to the Carpathian Basin, the Balkans, through Greece and back into Anatolia and the Near East where farming originated ~12,000 years ago[56,57]. For the three subsequent Early Neolithic cultures with a smaller geographical distribution, a cultural origin outside the MES region (but nevertheless in Central Europe) is assumed, which explains the temporal overlap of the beginning and ending of subsequent cultures:
Linienbandkeramik (LBK): 5450–4775 BC,
Stichbandkeramik (SBK): 4925–4625 BC,
Rössen (RSK): 4625–4475/4250 BC and

Gatersleben (GLK): 4475–4250 BC.

The LBK is the oldest Neolithic culture in Central Germany and represents the first people in the region who relied on agriculture and animal husbandry. The LBK was identified earliest in Western Hungary (Transdanubia), where it incorporated novel technologies and ideas from Anatolia, the Levant and the Near East during the Neolithic transition. From there, the LBK expanded relatively rapidly along the major waterways and fertile Loess towards Central Europe, reaching as far as the Paris Basin in the West and Ukraine in the East[58]. The terminal Early Neolithic is characterised by *Epi-Rössen* and *Epi-Lengyel*, two horizons (Late-*Lengyel*-horizon; 4250-3950 BC) that are defined by distinct ceramics found at the sites *Gröna* and *Schiepzig*, respectively.

Middle Neolithic:

The Middle Neolithic (4250–2725 BC) in Saxony-Anhalt encompasses eight cultures: *Michelsberg, Baalberge*, *Salzmünde*, *Tiefstichkeramik*, *Walternienburg*, *Bernburg*, *Elb-Havel* and *Kugelamphorenkultur* (Globular Amphorae). Most of these cultures are only present in distinct regions of Saxony-Anhalt and have emerged locally, whereas others originated outside of it. This explains the large overlap of temporal successions and partial contemporaneity/co-existence in neighbouring regions:
Michelsberg (MBK): 4250–3500 BC
Jordansmühle (JMK): 4100–3650 BC
Baalberge (BAC): 3950–3400 BC
Salzmünde (SMC): 3400–3100/3025 BC
Tiefstichkeramik (TSC): 3650–3325 BC
Walternienburg (WBC): 3325–3100 BC
Bernburg (BEC): 3100–2650 BC
Elb-Havel (EHC): 3100–2725 BC

By and large, the Mid-Neolithic in Saxony-Anhalt is defined by two phases of successive cultures unified by the overarching theme of the Funnel beaker tradition, an older phase (3950–3100 BC) and younger one (3100–2650 BC). Some of these cultures are interlocked with previous ones, whereas others seem to replace existing ones, arguing for differential processes over time and space. Towards the end of the Mid-Neolithic, a seemingly foreign culture (*Fischbeck* group) whose origin is not yet resolved can be observed in eastern parts of the Mittelelbe-Saale area, but which resembles the late Globular Amphorae culture, and gives rise to the *Schönfeld* culture.

Late Neolithic:

The Late Neolithic in MES is represented by five cultures, including the
Schnurkeramik or *Corded Ware* culture (CWC): 2800–2200/2050 BC,
Einzelgrab or *Single-grave* culture (SGC): 2800–2200/2050 BC,
Schönfeld culture in the Northeast (SFC): 2725–2050 BC,
Ammensleben: 2650–2200/2050 BC, and
Glockenbecher or *Bell Beaker* culture/phenomenon (BBC): 2500–2200/2050 BC.

The Late Neolithic horizon is defined as 2725–2200 BC, even though earliest signs of the Corded Ware culture can be found around 2800 BC, whereas remains of beaker cultures can last as long as 2050 BC. The Bell Beaker culture is evident in the archaeological record from 2500 BC onwards in the South of MES, when it starts to move into settlement areas previously occupied by Corded Ware people, the latter of which have affinities to archaeological groups further east[36]. The settlement density in MES increases during its later phase (2300-2050 BC), when the Corded Ware is superseded by Bell Beaker elements, often at the same sites. This Late Neolithic Bell Beaker phenomenon is of particular interest, since archaeological evidence suggests it originated in the Tagus region of Western Iberia around 2800-2700 BC before spreading to become one of the first major pan-European cultures[59]. It has been traced archaeologically over Western Europe (including enclaves in North Africa) as far as Hungary, Ireland, and southern Scandinavia. Earlier Neolithic cultures were overlain/infiltrated by discernable Bell Beaker elites with a cultural package endowed with rich grave goods (including the eponymous bell-shaped ceramic beakers).

During the transition to the Bronze Age, early Bronze Age cultural elements of the Unetice culture appear contemporaneously to late elements of the Bell Beaker culture, again sometimes also at the same site. The site of Eulau, famously known for its oldest nuclear family graves, represents a good example of cultural dynamics during the Late Neolithic, as it shows the presence of both Corded Ware and Bell Beaker cultures, and later of the early Bronze Age Unetice culture.

6. Additional discussion of genetic results in the light of archaeological and anthropological evidence

The expansion of hg H is compatible with a post-glacial resettling of Northern and Central Europe from southwestern refugia[60-62], as indicated by our population skyride and PCA plot. However, we propose that the rise of hg H to become the predominant mtDNA branch in Europe was mediated by subsequent demographic events during the Neolithic, as evidenced by a general increase in hg H frequency and strong population growth during this period. Interestingly, hg H was virtually absent amongst Central and Northern European

hunter-gatherers[63,64], formed only 19% of LBK individuals, ~20% of a French Middle Neolithic (3030–2890 calBC) site[65] and ~28% of Neolithic samples from the Iberian Peninsula[66,67]. The widespread Bell Beaker phenomenon fits well as a potential driving factor for the dissemination of hg H throughout most parts of Western Europe. Having reached a certain frequency, and assuming a generally higher rate of population growth in southern and western Europe in post-Neolithic times[68], this may have contributed to the fixation of H as the predominant European mtDNA haplogroup.

The spread of both the LBK and Bell Beaker cultures across large parts of Europe might be considered exceptional, supra-regional phenomena. Assuming that the observed match between ancient hg H sub-clade frequencies and particular present-day population groupings does reflect ancient 'genetic homelands', Central European LBK and Bell Beaker populations must have carried strong signals of a geographical and cultural origin in their (mt) genome, here in the form of characteristic subsets of hg H sub-clades. The fact that these signals were not lost (or obscured by later events) suggests two cultures with considerable coherence as interbreeding human populations over time (many generations) and space (from the Near East or the Iberian peninsula to Central Europe). Craniometric and morphological data from Mesolithic hunter-gatherer and Neolithic farming populations also suggest a significant population influx from the Near East into Europe[69] and substantial variability throughout the Neolithic[70]. Distinct mobility profiles have also been identified via strontium isotopes for both the LBK[71] and Bell Beaker period[72].

However, in all likelihood, most cultural change observable in the Central European archaeological record probably did not involve significant populations of highly mobile individuals driven by new technologies, economic systems, and/or cultural beliefs[73] to cover wide and disparate parts of Europe (and thereby leave an identifiable genetic mark). Therefore, we might not expect distinct genetic signatures from all archaeological cultures, particularly not from more regionally restricted ones.

In general, while our current estimates of genetic affinities are somewhat limited by relatively small sample numbers, a more detailed picture will undoubtedly emerge as more data become available. However, our use of complete mt genomes to perform a diachronic study showcases the enormous potential for fine-scale reconstructions of Europe's prehistory.

Supplementary Table 11. Detailed archaeological information for each individual.

Individual	Site	Culture	Radiocarbon dates, uncalibrated	Collection details				
				Museum no.	feature	find	grave	area
HAL36	Halberstadt-Sonntagsfeld	Linear Pottery culture LBK (5450-4775 BC)		2000:4338a	1114		40	05
HAL11	Halberstadt-Sonntagsfeld			2000:3988a	340		9	02
HAL32	Halberstadt-Sonntagsfeld			2000:4307a	1059		36	04
HAL39	Halberstadt-Sonntagsfeld		6145 ± 30 BP (KIA40343)	2000:4014a	413.1		11	01
DEB9	Derenburg-Meerenstieg II			1998:1175a	420		9	5896
DEB21	Derenburg-Meerenstieg II		6,151+/-27 BP (KIA30404)	1998:1288a	600		32	5897
KAR6a	Karsdorf			2006:14423a	170			2006
KAR11b	Karsdorf			2004:26267a	430			2004
KAR16a	Karsdorf			2004:26374a	611			2004
OSH2	Oberwiederstedt-	Rössen (4625-4475/4250 BC)			195			
OSH3	Oberwiederstedt-Schrammhöhe				206			
OSH1	Oberwiederstedt-Schrammhöhe				225			
OSH7	Oberwiederstedt-Schrammhöhe		5669 ± 54 (Erl-8395)		95			
SALZ18a	Salzmünde	Schöningen (4100-3950 BC)		2007:7446	3932	2		
SALZ21b	Salzmünde			2006:4805	4090	1		
ESP30	Esperstedt	Baalberge (3950-3400 BC)	5061 ± 62 (Er17784)	2004:22538	6220			A38
HQU4	Halle-Queis			2002:2328a	957			
SALZ57a	Salzmünde	Salzmünde (3400-3100/3025 BC)	4498 ± 27 (KIA-31459)	2006:5445	3833	1		
SALZ77a	Salzmünde			2006:6405	5533	2		
ESP15	Esperstedt	Corded Ware (2800-2200/2050 BC)	3904 ± 47 (Er17257)	2004:21022	6			A38
BZH6	Benzingerode-Heimburg			2003:2314	1287	1036	2	
BZH4	Benzingerode-Heimburg	Bell Beaker (2500-2200/2050 BC)		2003:2589	4607	2267	7	
ROT6	Rothenschirmbach		3953 ± 47 (Erl 8710)	2005:1685	10044			A38
ALB1	Alberstedt		3858 ± 57 (Er18537)	2005:3369	7136			A38
ROT1	Rothenschirmbach		3818 ± 48 (Erl 8715)	2005:4167	10294			A38
ROT2	Rothenschirmbach			2005:4168	10293			A38
QUEXII1	Quedlinburg XII		3792 ± 50 (Erl-7042)	2004:11864a	1266.1			12

Individual	Site	Culture	Radiocarbon dates, uncalibrated	Collection details				
				Museum no.	feature	find	grave	area
QUEXII2	Quedlinburg XII		3655 ± 48 (Erl-7041)	2004:11857	1265			12
QLB26a	Quedlinburg		3839 ± 55 (Erl-8558)	2004:40935	19614	6a		7.1
QUEXII3	Quedlinburg XII		3820 ± 42 (Erl-7038)	2004:9470a	6256			12
QLB28b	Quedlinburg			2004:40957	19617	9a		7.1
BZH1	Benzingerode-Heimburg	Unetice (2200-1575 BC)		2003:2641	6035	3479	23	
BZH8	Benzingerode-Heimburg		3626 ± 27 (KIA27958)	2003:2650	5236.1	2380	12	
BZH14	Benzingerode-Heimburg		3511 ± 29 (KIA27953)	2003:2662	3109.2	3102b	24	
EUL41a	Eulau				882	13a		4
EUL57b	Eulau				1911.5	312, XA		5
QUEVIII4	Quedlinburg VIII		3587 ± 55 (Erl-7045)	2004:10747a	3646	5a		8
Sardinia	Si Bittuleris, Sardinia, Italy	Nuragic Bronze Age (1624 BC)	3398±26 (OxA-22193), associated	732				
Latsch	Latsch, South Tyrol, Italy	Iron Age (~500 BC)	2567± 45 (LTL2778A), associated	305, 103 III Tag				

References

1 Brandt, G., Knipper, C., Roth, C., Siebert, A. & Alt, K. W. in *Anthropologie, Isotopie und DNA, 2. Mitteldeutscher Archäologentag.*3 edn (eds H. Meller & K. W. Alt) 17-32 (Landesmuseum for Vorgeschichte Halle (Saale)).

2 Haak, W. *et al.* Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science (New York, N.Y.)* **310**, 1016-1018, doi:10.1126/science.1118725 (2005).

3 Haak, W. *et al.* Ancient DNA from European Early Neolithic Farmers Reveals Their Near Eastern Affinities. *PLoS Biology* **8**, e1000536, doi:10.1371/journal.pbio.1000536 (2010).

4 Taylor, R. W., Taylor, G. A., Durham, S. E. & Turnbull, D. M. The determination of complete human mitochondrial DNA sequences in single cells: implications for the study of somatic mitochondrial DNA point mutations. *Nucleic Acids Research* **29**, E74-74 (2001).

5 Andrews, R. M. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat.Genet.* **23**, 147 (1999).

6 Patel, M. & Sive, H. PCR-based subtractive cDNA cloning. *Current protocols in molecular biology* **Chapter 25**, Unit 25B 22, doi:10.1002/0471142727.mb25b02s55 (2001).

7 Tao, S. C., Gao, H. F., Cao, F., Ma, X. M. & Cheng, J. Blocking oligo--a novel approach for improving chip-based DNA hybridization efficiency. *Molecular and cellular probes* **17**, 197-202 (2003).

8 Brotherton, P. *et al.* 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**, 5717-5728, doi:10.1093/nar/gkm588 (2007).

9 Zhou, S. *et al.* An oligonucleotide microarray for high-throughput sequencing of the mitochondrial genome. *The Journal of molecular diagnostics : JMD* **8**, 476-482, doi:10.2353/jmoldx.2006.060008 (2006).

10 Paabo, S. *et al.* Genetic analyses from ancient DNA. *Annu Rev Genet* **38**, 645-679, doi:10.1146/annurev.genet.37.110801.143214 (2004).

11 Brotherton, P., Sanchez, J. J., Cooper, A. & Endicott, P. Preferential access to genetic information from endogenous hominin ancient DNA and accurate quantitative SNP-typing via SPEX. *Nucleic Acids Research* **38**, e7, doi:10.1093/nar/gkp897 (2010).

12 Malmström, H. *et al.* More on contamination: the use of asymmetric molecular behavior to identify authentic ancient human DNA. *Mol Biol Evol* **24**, 998-1004, doi:10.1093/molbev/msm015 (2007).

13 Adler, C. J., Haak, W., Donlon, D. & Cooper, a. Survival and recovery of DNA from ancient teeth and bones. *Journal of Archaeological Science* **38**, 956-964, doi:10.1016/j.jas.2010.11.010 (2011).

14 Maitra, A. *et al.* The Human MitoChip: a high-throughput sequencing microarray for mitochondrial mutation detection. *Genome Research* **14**, 812-819, doi:10.1101/gr.2228504 (2004).

15 Mithani, S. K. *et al.* Mitochondrial resequencing arrays detect tumor-specific mutations in salivary rinses of patients with head and neck cancer. *Clin Cancer Res* **13**, 7335-7340, doi:10.1158/1078-0432.CCR-07-0220 (2007).

16 Hartmann, A. *et al.* Validation of Microarray-Based Resequencing of 93 Worldwide Mitochondrial Genomes. *Human Mutation* **30**, 115-122, doi:Doi 10.1002/Humu.20816 (2009).

17 *Affymetrix GeneChip CustomSeq Resequencing Array Protocol*, <http://www.affymetrix.com/support/mas/index.affx - 1_2> (2011).

18 Walsh, D. J. *et al.* Isolation of deoxyribonucleic acid (DNA) from saliva and forensic science samples containing saliva. *Journal of forensic sciences* **37**, 387-395 (1992).

19 Green, R. E. *et al.* Analysis of one million base pairs of Neanderthal DNA. *Nature* **444**, 330-336, doi:10.1038/nature05336 (2006).
20 Bruce, C. K. *et al.* Design and validation of a metabolic disorder resequencing microarray (BRUM1). *Human Mutation* **31**, 858-865, doi:10.1002/humu.21261 (2010).
21 Hofreiter, M., Jaenicke, V., Serre, D., von Haeseler, A. & Paabo, S. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Research* **29**, 4793-4799 (2001).
22 Hansen, A. J., Willerslev, E., Wiuf, C., Mourier, T. & Arctander, P. Statistical Evidence for Miscoding Lesions in Ancient DNA Templates. *Mol.Biol.Evol.* **18**, 262-265 (2001).
23 Briggs, A. W. *et al.* Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci U S A* **104**, 14616-14621, doi:10.1073/pnas.0704665104 (2007).
24 Haak, W. *et al.* Ancient DNA, Strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age. *Proc Natl Acad Sci U S A* **105**, 18226-18231, doi:10.1073/pnas.0807592105 (2008).
25 Anderson, S. *et al.* Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457-465 (1981).
26 van Oven, M. & Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human Mutation* **30**, E386-394, doi:10.1002/humu.20921 (2009).
27 Bandelt, H. J., Kong, Q. P., Richards, M. & Macaulay, V. in *Human Mitochondrial DNA and the Evolution of Homo Sapiens* (eds H. J. Bandelt, V. Macaulay, & M. Richards) 47-90 (Springer, 2006).
28 Behar, D. M. *et al.* The Genographic Project public participation mitochondrial DNA database. *PLoS Genet* **3**, e104, doi:10.1371/journal.pgen.0030104 (2007).
29 Strachan, T. & Read, A. P. *Human Molecular Genetics*. 2nd edn, (Wiley-Liss, 1999).
30 Xie, H. M. *et al.* Mitochondrial genome sequence analysis: A custom bioinformatics pipeline substantially improves Affymetrix MitoChip v2.0 call rate and accuracy. *BMC Bioinformatics* **12**, 402, doi:10.1186/1471-2105-12-402 (2011).
31 Travers, K. J., Chin, C. S., Rank, D. R., Eid, J. S. & Turner, S. W. A flexible and efficient template format for circular consensus sequencing and SNP detection. *Nucleic Acids Research* **38**, e159, doi:10.1093/nar/gkq543 (2010).
32 Geneious v5.4, Available from http://www.geneious.com/ (2011).
33 Behar, D. M. *et al.* A "Copernican" Reassessment of the Human Mitochondrial DNA Tree from its Root. *American Journal of Human Genetics* **90**, 1-10, doi:10.1016/j.ajhg.2012.03.002 (2012).
34 R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2010).
35 Nowak, M. Transformations in East-Central Europe from 6000 to 3000 BC: local vs. foreign patterns. *Documenta Praehistorica* **XXXIII**, 143-158 (2006).
36 Bogucki, P. I. & Crabtree, P. J. *Ancient Europe 8000 B.C.--A.D. 1000 : Encyclopedia of the Barbarian World*. (Charles Scribner's Sons, 2004).
37 Alvarez-Iglesias, V. *et al.* New population and phylogenetic features of the internal variation within mitochondrial DNA macro-haplogroup R0. *PLoS ONE* **4**, e5112, doi:10.1371/journal.pone.0005112 (2009).
38 Garcia, O. *et al.* Using mitochondrial DNA to test the hypothesis of a European post-glacial human recolonization from the Franco-Cantabrian refuge. *Heredity* **106**, 37-45, doi:10.1038/hdy.2010.47 (2011).
39 Roostalu, U. *et al.* Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: The near eastern and Caucasian perspective. *Molecular Biology and Evolution* **24**, 436-448 (2007).
40 Brandstätter, A. *et al.* Dissection of mitochondrial superhaplogroup H using coding region SNPs. *Electrophoresis* **27**, 2541-2550, doi:10.1002/elps.200500772 (2006).

41 Loogväli, E. L. *et al.* Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. *Mol.Biol.Evol.* **21**, 2012-2021 (2004).
42 Hammer, O., Harper, D. A. T. & Ryan, P. D. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica* **4** (2001).
43 Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567, doi:10.1111/j.1755-0998.2010.02847.x (2010).
44 Posada, D. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* **25**, 1253-1256, doi:10.1093/molbev/msn083 (2008).
45 Schneider, S. & Excoffier, L. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* **152**, 1079-1089 (1999).
46 Excoffier, L. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology* **13**, 853-864, doi:10.1046/j.1365-294X.2003.02004.x (2004).
47 Behar, D. M. *et al.* The Basque Paradigm: Genetic Evidence of a Maternal Continuity in the Franco-Cantabrian Region since Pre-Neolithic Times. *American Journal of Human Genetics* **90**, 486-493, doi:10.1016/j.ajhg.2012.01.002 (2012).
48 Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**, 221-224, doi:10.1093/molbev/msp259 (2010).
49 Keane, T. M., Creevey, C. J., Pentony, M. M., Naughton, T. J. & McLnerney, J. O. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol Biol* **6**, 29, doi:10.1186/1471-2148-6-29 (2006).
50 Minin, V. N., Bloomquist, E. W. & Suchard, M. A. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol Biol Evol* **25**, 1459-1471, doi:10.1093/molbev/msn090 (2008).
51 Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214, doi:10.1186/1471-2148-7-214 (2007).
52 Rambaut, A. & Drummond, A. *Tracer v1.4, Available from* http://beast.bio.ed.ac.uk/Tracer, 2007).
53 Fu, Q., Rudan, P., Pääbo, S. & Krause, J. Complete Mitochondrial Genomes Reveal Neolithic Expansion into Europe. *PLoS ONE* **7**, e32473. doi:32410.31371/journal.pone.0032473 (2012).
54 Ho, S. Y. *et al.* Bayesian estimation of substitution rates from ancient DNA sequences with low information content. *Systematic biology* **60**, 366-375, doi:10.1093/sysbio/syq099 (2011).
55 Schwarz, R. in *Kataloge zur Dauerausstellung im Landesmuseum für Vorgeschichte Halle, 2. Neolithikum* (ed H. Meller) (in the press).
56 Price, T. D. *Europe's first farmers*. (Cambridge University Press, 2000).
57 Whittle, A. W. R. & Cummings, V. *Going over: the mesolithic-neolithic transition in North-West Europe*. (Oxford University Press, 2007).
58 Gronenborn, D. A variation on a basic theme: the transition to farming in southern central Europe. *J.World Prehistory* **13**, 123-210 (1999).
59 Case, H. in *Beyond Stonehenge. Essays in honour of Colin Burgess* (eds C. Burgess, P. Topping, & F. Leach) (2007).
60 Pereira, L. *et al.* High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. *Genome Res* **15**, 19-24, doi:15/1/19 10.1101/gr.3182305 (2005).
61 Torroni, A. *et al.* A signal, from human mtDNA, of postglacial recolonization in Europe. *American Journal of Human Genetics* **69**, 844-852 (2001).

62 Achilli, A. *et al.* The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am J Hum Genet* **75**, 910-918 (2004).
63 Bramanti, B. *et al.* Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science (New York, N.Y.)* **326**, 137-140, doi:10.1126/science.1176869 (2009).
64 Malmström, H. *et al.* Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. *Current biology : CB* **19**, 1758-1762, doi:10.1016/j.cub.2009.09.017 (2009).
65 Lacan, M. *et al.* Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. *P Natl Acad Sci USA* **108**, 9788-9791, doi:10.1073/pnas.1100723108 (2011).
66 Sampietro, M. L. *et al.* Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe. *Proceedings. Biological sciences / The Royal Society* **274**, 2161-2167, doi:10.1098/rspb.2007.0465 (2007).
67 Gamba, C. *et al.* Ancient DNA from an Early Neolithic Iberian population supports a pioneer colonization by first farmers. *Molecular Ecology* **21**, 45-56, doi:10.1111/j.1365-294X.2011.05361.x (2012).
68 Livi-Bacci, M. *A Consice History of World Population*. Fourth edn, (Blackwell Publishing, 2007).
69 Pinhasi, R. & von Cramon-Taubadel, N. Craniometric data supports demic diffusion model for the spread of agriculture into Europe. *PLoS ONE* **4**, e6747, doi:10.1371/journal.pone.0006747 (2009).
70 Gallagher, A., Gunther, M. M. & Bruchhaus, H. Population continuity , demic diffusion and Neolithic origins in central-southern Germany : The evidence from body proportions. *Human Biology* **60**, 95-126, doi:10.1016/j.jchb.2008.05.006 (2009).
71 Price, T. D., Bentley, R. A., Lüning, J., Gronenborn, D. & Wahl, J. Prehistoric human migration in the Linearbandkeramik of Central Europe. *Antiquity* **75**, 593-603 (2001).
72 Grupe, G., Price, T. D. & Sollner, F. Mobility of Bell Beaker people revealed by strontium isotope ratios of tooth and bone: a study of southern Bavarian skeletal remains. A reply to the comment by Peter Horn and Dieter Muller-Sohnius. *Applied Geochemistry* **14**, 271-275 (1999).
73 Heyd, V. Families, prestige goods, warriors & complex societies: Beaker groups of the 3rd millennium cal BC along the upper & middle Danube. *P Prehist Soc* **73**, 327-379 (2007).

Chapter 6

Ancient whole mitochondrial genomes to study recent human evolution

(Part 2)
Re-calibrating the timescale of recent human evolution using whole mitochondrial genomes from the Neolithic

Statement of authorship

Re-calibrating the timescale of recent human evolution using whole mitochondrial genomes from the Neolithic

Julien Soubrier (Candidate)
Co-developed the research concept, performed the phylogenetic analyses, interpreted the results, developed the figures and wrote the manuscript.
I hereby certify that the statement of contribution is accurate.

Signed .... *Date* 02.04.12

Simon Y. W. Ho
Co-developed the research concept, co-interpreted the results, supervised phylogenetic analyses and manuscript preparation.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .... *Date* .....05.04.2012.....

Paul Brotherton
Devised the library preparation and target enrichment protocol.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed ..... *Date* 8/4/12

Wolfgang Haak
Co-developed the research concept, processed the samples, analysed data, and supervised manuscript preparation.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .. .. *Date* 02/04/2012

Alan Cooper
Co-developed the research concept, and co-wrote the manuscript.
I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signed .. .. *Date* .....

Re-calibrating the timescale of recent human evolution using whole mitochondrial genomes from the Neolithic

Julien Soubrier[1], Simon Y. W. Ho[2], Paul Brotherton[1§], Wolfgang Haak[1], Alan Cooper[1] and The Genographic Consortium

[1] Australian Centre for Ancient DNA, School of Earth & Environmental Sciences, University of Adelaide, Adelaide, SA, Australia

[2] School of Biological Sciences, University of Sydney, Sydney, NSW, Australia

[§] New address: Archaeogenetics Research Group, School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, UK.

ABSTRACT

Anatomically modern humans appeared in Africa between 100,000 and 200,000 years ago, but most parts of the world have only been colonised within the past 50,000 years. Phylogenetic information from the mitochondrial genome and other genetic markers has contributed to the global picture of the human evolutionary history. However, the precise timescale of this recent human evolution and dispersal is still subject to debate. Most estimates of this timescale are calibrated using the human-chimpanzee divergence, estimated at around 6.5 My from the fossil record. However, the considerable disparity between estimates based on this fossil calibration and those based on more recent calibration points, such as biogeographical events or known pedigrees, has raised doubts about the reliability of current inferences.

In the present study we use 46 whole mitochondrial genomes from dated ancient human samples to re-calibrate the timescale of recent human evolution. We show that the dates associated with the ancient samples are sufficient to calibrate the entire phylogeny of modern humans. Our estimates of mitochondrial molecular rates are higher than those from previous studies based on fossil calibration, but lower than the estimates from pedigree studies. This result is in accordance with the time dependence of molecular rates and has important consequences for the estimated timescale of human dispersals across the globe.

INTRODUCTION

From our origins in Africa, the global migration routes taken by anatomically modern humans (AMH) are now well established. However, the exact timescale of this journey is still subject to debate (Stoneking et al. 1992; Endicott et al. 2009; Oppenheimer 2012). This temporal uncertainty is mainly due to the lack of appropriate calibration points covering the past 200 thousand years (ky) of human evolution. Most analyses of human evolutionary timescales have been estimated using the human-chimpanzee split as an external calibration point (e.g., Mishmar et al. 2003; Soares et al. 2009). However, molecular rates depend on the time period over which they are inferred (see Chapter 5: Ho et al. 2011a). More precisely, the depth of a particular calibration point has an impact on the estimated rates and dates. For example, coalescent dates of shallow timescales tend to be overestimated when they are based on old fossil calibrations (examples are given in Chapter 2). Considering this time dependence of molecular rates, the split between human and chimpanzee, at 6.5 million years before present (My BP), is too old to be able to provide a precise calibration for the evolutionary history of modern humans (Stoneking et al. 1992; Ho et al. 2005; Henn et al. 2009; Subramanian et al. 2009; Ho et al. 2011a). In contrast, mutation rates estimated in pedigree studies tend to be very high (Howell et al. 2003; Santos et al. 2005; Madrigal et al. 2012), as they account for genetic changes that will not necessarily be fixed in the population. The use of a high rate estimate will lead to coalescent age underestimates. It appears that only calibration points from the same time period than the one targeted would provide an accurate estimate of the timescale (Ho and Larson 2006).

So far, the main solution to overcome the lack of direct calibration along the human tree has been the use of archaeological evidence for biogeographic events, such as the colonisation of islands or continents (Stoneking et al. 1992; Atkinson, Gray, Drummond 2008; Endicott, Ho 2008; Henn et al. 2009). However, the link between the divergences of genetic lineages and populations remains uncertain, as genetic isolation between two populations could have happened long before they became geographically separate. Alternatively, gene flow could have occurred or continued between two populations after and despite their geographic separation.

The use of ancient DNA offers a useful alternative to calibrate estimates of short evolutionary timescales. In this case, radiocarbon dates associated with ancient sequences can provide direct temporal information at the tips of the evolutionary tree (Drummond et al. 2003; Ho et al. 2011b). This removes the need to make assumptions about the correspondence between genetic and population divergences, or about the consistency of rates between extra- and intraspecific levels (see Chapter 5). In this context, our capacity to infer a precise timescale from ancient AMH sequences will depend on their distribution across the human

tree, both over time and among lineages (Ho et al. 2011b). Published studies of ancient human samples have involved either very short fragments (HVR) or only few whole mitochondrial genomes (WMG) (Ötzi/Tyrolean Iceman, Ermini et al. 2008; Saqqaq, Gilbert et al. 2008; Kostenki, Krause et al. 2010a), providing inadequate temporal information to calibrate estimates of the timescale of human evolution and dispersal. In the present study, we use 43 novel WMG sequences from ancient AMH samples, in addition to the existing sequences of AMH (3 WMG), Neandertal (5 WMG) and Denisovan (1 WMG), to re-calibrate the human mitochondrial timescale.

In contrast to studies using only short fragments of the control region, the use of whole mitochondrial genomes drastically increases the amount of phylogenetic information. However, the heterogeneity of evolutionary processes in different parts of the mitochondrial genome needs to be taken into account. For example, the control region is known to have a higher mutation rate than the coding region. This rate heterogeneity can potentially lead to biased estimates of rates and dates, as shown in Chapter 5. Therefore, partitioning the WMG into different subsets, with a different molecular clock model for each of them, is required to obtain reliable estimates of rates and dates. Most of the partitioning schemes commonly used for the mitochondrial genome are based on subjective rules (e.g., one subset per gene or per codon position, or different subsets for non-coding and coding regions), without statistical justification. To objectively define the best partitioning scheme for each sequence alignment, we use the program PartitionFinder (Lanfear et al. 2012).

Refinement of the human mitochondrial timescale can help to clarify the uncertainties that are still persistent in the evolutionary history of AMH. For example, while the presence of AMH outside Africa before the Toba eruption (which represents an important palaeontological stratum at 74 ky BP) is unconfirmed by the archaeological record, the presence of closely related hominids in South-East Asia before that period lead to confounding palaeoanthropological evidence on the subject (Oppenheimer 2012). Also, the potential role of climate change in the expansion of human population could be more accurately assessed from precisely timed demographic events (Atkinson, Gray, Drummond 2008; Stewart, Stringer 2012).

MATERIAL AND METHOD

Data

Ancient sequences

As part of The Genographic Project, complete mitochondrial genomes were sequenced from 43 ancient human remains at the Australian Centre for Ancient DNA (Table 1). The method used to extract, target-enrich, and sequence WMG is provided in the first part of Chapter 6, along with population genetic analyses of the 37 hg H sequences. The additional six sequences belong to the haplogroup (hg) N1a. In addition to these newly sequenced WMGs, nine published sequences from AMH, Neandertal and Denisovan samples were included in the study (Ermini et al. 2008; Gilbert et al. 2008; Green et al. 2008; Briggs et al. 2009; Krause et al. 2010a; Krause et al. 2010b):

Modern sequences

To obtain data sets of reasonable size for Bayesian analysis, two sets of 154 and one set of 254 AMH sequences were randomly selected from a pool of 1,127 non-redundant WMGs (Phylotree, van Oven, Kayser 2009, http://www.phylotree.org, mtDNA tree Build 4). We added data from 46 ancient individuals to obtain data sets of 200 and 300 AMH sequences each (Table 2). In addition, one alignment of 206 hominids was obtained by adding six sequences from Neandertal and Denisovan individuals to a set of 200 AMH. Independently, one N1a dataset was built by combining the 20 N1a modern mitochondrial sequences available on Phylotree (Build 13, Dec 28 2011, Palanichamy et al. 2010) with the six newly sequenced WMGs from ancient samples.

Table 1. List of all ancient samples used in the study

	Sub-species	Sample	hg	Age (years)
ACAD	AMH hg H	ALB1	H3b	4,276
		BZH1	H11a	4,225
		BZH14	H82a	3,834
		BZH4	H1e8	4,225
		BZH6	H1ca1	4,375
		BZH8	H2a1a3	3,959
		DEB21	H1j	7,072
		DEB9	H88	7,060
		ESP15	H6a1a	4,380
		ESP30	H1e1a5	5,792
		EUL41a	H4a1a1a2	3,837
		EUL57b	H3	3,837
		HAL11	H	7,060
		HAL32	H26	7,060
		HAL36	H23	7,060
		HAL39	H1e	7,134
		HQU4	H7d5	5,625
		KAR11b	H	7,060
		KAR16a	H46b	7,060
		KAR6a	H1bz	7,060
		Latsch	H16	2,450
		OSH1	H89	6,350
		OSH2	H1	6,350
		OSH3	H5b	6,350
		OSH7	H1	6,476
		QLB26a	H1	4,310
		QUEVIII4	H7h	3,980
		QUEXII1	H4a1	4,250
		QUEXII2	H4a1	4,000
		QUEXII3	H13a1a2c	4,290
		ROT1	H3ao2	4,221
		ROT6	H5a3	4,416
		SALZ18a	H10i	5,975
		SALZ21b	H1e7	5,975
		SALZ57	H3	5,248
		SALZ77a	H3	5,181
		Sardinia	H1aw1	3,574
	AMH hg N1a	OAO1	N1a1	6,387
		HAL4	N1a1	7,062
		HAL7	N1a1b	7,062
		HAL15	N1a1a3	7,062
		HAL27	N1a1a3	7,062
		HAL34	N1a1b	7,062
Published	AMH	EU810403	K1 (Ötzi)	5,200
		EU725621	D2a1 (Saqqaq)	4,000
		FN600416	U2 (Kostenki)	30,000
	Denisovan	NC_013993	-	35,000
	Neandertal	NC_011137	-	38,310
		FM865407	-	39,900
		FM865408	-	39,240
		FM865409	-	38,790
		FM865411	-	65,000

We also had access to novel hg H samples via collaborators Doron M. Behar and Lluis Quintana-Murci, providing 729 non-redundant hg H WMGs out of the 1,236 samples. These were recently published (Behar et al. 2012a, 2012b) and are listed in the updated built of PhyloTree (Build 14, Apr 5 2012). Among the 729 H WMGs, one set of 200 and one set of 300 sequences were randomly selected and combined with the 37 hg H ancient WMGs (Table 2).

In order to replicate the analyses based on the human/chimpanzee tMRCA (time of the most common recent ancestor) calibration, we also collected five chimpanzee WMGs and combined them with one set of 154 modern AMH sequences.

Table 2. Summary of the eight data sets used in the present study.

	AMH	**AMH**	**AMH**	**AMH/Neand/ Denis**	**hg H**	**hg H**	**hg N1a**	**AMH/ Chimp**
Samples total	**200 (1)**	**200 (2)**	**300**	**206**	**237**	**337**	**26**	**159**
No. modern AMH samples, random	154	154	254	154	200	300	20	154
No. ancient AMH samples	46	46	46	46	37	37	6	-
Denisovan	-	-	-	1	-	-	-	-
Neandertal	-	-	-	5	-	-	-	-

Analyses

Sequence alignment

Sequences were aligned to the revised Cambridge Reference Sequence (rCRS: AC_000021) using the program SeaView (Gouy, Guindon, Gascuel 2010) and adjusted manually. We excluded transversions and insertions at np 309.1C(C), 315.1C, 523-524d (aka 522-523d), 16182C, 16183C, 16193.1C(C) and a transition at np 16519 from phylogenetic reconstruction as suggested in previous studies (e.g., Bandelt et al. 2006; van Oven, Kayser 2009).

Partitioning

The partitioning scheme and the substitution model for each subset were simultaneously selected using PartitionFinder (Lanfear et al. 2012) using the Bayesian information criterion. Considering all possible partitioning schemes from nine predefined subsets (all three codon positions of the coding genes and ND6 subdivided into CP1 (first codon position), CP2, CP3, ND6 CP1, ND6 CP2, ND6 CP3, as well as tRNA genes, rRNA genes, and the D-loop), the best scheme identified by PartitionFinder was a division in five subsets, combining the three ND6 codon positions and CP1, tRNA and rRNA as one partition each. However, since there is no *a priori* biological support for the combination of CP1 and t/rRNA, and to allow for comparison of our results with previous studies, we set Coding CP1

and t/rRNA apart as two subsets. This lead to the following scheme of six partitions: coding genes CP1, CP2, CP3, as well as ND6, tRNA + rRNA genes, and the D-loop.

To account for nine insertions observed in the phylogeny of hg H (Phylotree) we added 5 positions to the D-loop subset (42.1G; 44.1C; 291.1A; 573.1CC), and 4 to the RNA gene subset (5899.1C; 12310.1C; 960.1C; 5752.1A).

Bayesian phylogenetic analyses

Sequence alignments were analysed using BEAST 1.6.2 (Drummond, Rambaut 2007). For each subset the substitution model was selected through comparison of the Bayesian information criterion in PartitionFinder. An uncorrelated lognormal relaxed clock was used for each subset to account for potential rate variations between different branches of the phylogenetic tree (Drummond et al., 2006). A strict clock was used when it could not be rejected for a particular subset. The Bayesian skyride demographic model (Minin, Bloomquist, Suchard 2008) was used to account for demographic changes through time, and to reconstruct the demographic history of AMH.

As the skyride option is based on the coalescent model, it assumes a random sampling of individuals from a panmictic population. This assumption is violated when individuals of different species were used in the same analysis. We tested the impact of a violation of the panmictic model in a second analysis using the *BEAST option and chimpanzee samples to calibrate the human phylogeny. The *BEAST option combines the implementation of a coalescent model between individuals defined as part of the same population, and the Yule model at the inter-specific level, where each node represents a speciation event.

Date–randomisation test

A date-randomisation test was conducted on one alignment of 200 AMH sequences, one alignment of 237 hg H and the alignment of 26 hg N1a (Table 2) to test if the number of radiocarbon dates from ancient samples and its temporal spread were sufficiently large to calibrate estimates of evolutionary timescales (Ho et al. 2011b). Molecular rate estimates from the analysis with the true sample dates and five replicates with randomized dates were compared for each of the six mitochondrial subsets.

RESULTS AND DISCUSSION

Date-randomisation test

Figure 1 shows the comparison of estimated rates from the main analysis (true sample dates) and from replicates with randomised dates. For both the AMH and the hg H trees, the mean rate estimated for the six mitochondrial subsets is never included in any of the 95% highest posterior density intervals of the rate estimates obtained with randomised sample dates. This indicates that the dates of the ancient samples included in the alignments provide enough temporal information to calibrate the phylogenies from each dataset. However, for the N1a phylogeny, only the D-loop subset passes the date-randomisation test, revealing that the dates associated with the six ancient N1a sequences are not sufficient to calibrate an estimate of the rate. In some instances, the lack of temporal coverage from the ancient samples might be responsible for such a result, but in the case of hg N1a, the dates of the six ancient sequences range up to 7,000 years BP, similar to the dates used for hg H. Considering the low number of total samples for hg N1a and the low number of ancient sequences, it is likely that the observed lack of temporal signal comes from a lack of genetic diversity between modern and ancient sequences. If the ancient sequences represent haplotypes similar to those of modern sequences, there has been little genetic change between the sampling periods, and the rate estimate will be very imprecise. Consequently, further results are only presented for the hg H and the AMH phylogenies.

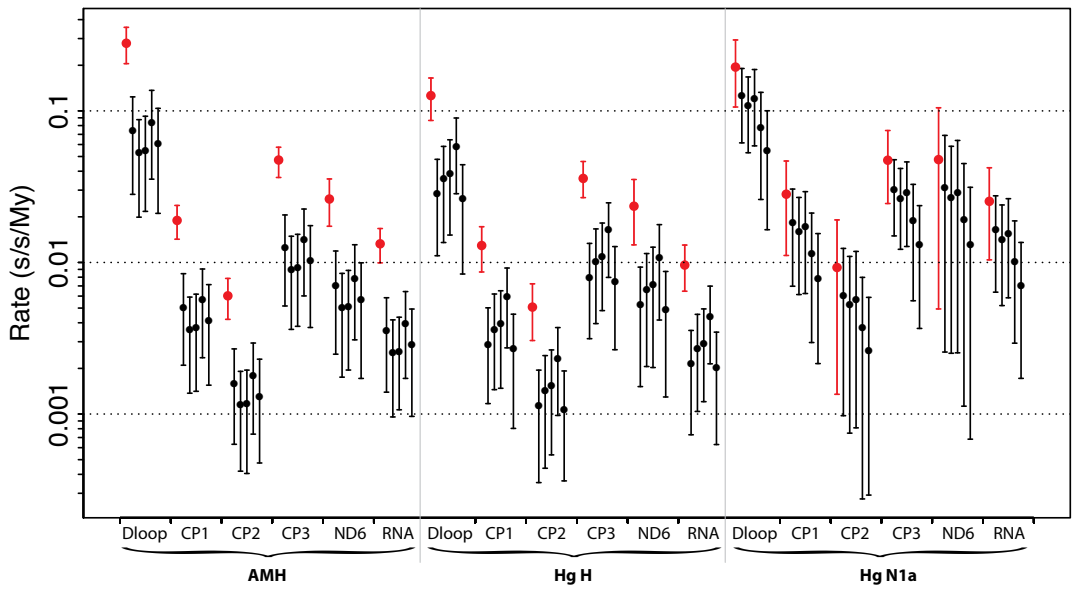


Figure 1. Results from date-randomisation tests for anatomically modern human (AMH), haplogroup H and haplogroup N1a data sets, with details for each subset. The red dots represent molecular rate estimates from the phylogenetic analyses calibrated using the radiocarbon dates associated with the ancient sequences. The black dots show the rates estimated with randomised dates. Lines represent the 95% highest posterior density of each rate estimate. The fact that none of the margins overlap with the original mean estimated rates of the AMH and haplogroup H demonstrates that the radiocarbon dates are informative enough to calibrate these rate estimates. For the N1a dataset, however, only the D-loop passes the test, indicating that the distribution of radiocarbon dates in time and among the genetic diversity of the clade is insufficient to calibrate rate estimates from the corresponding data set.

Estimates of rates and dates

Comparison of rate estimates between data sets

Estimated rates for the six mitochondrial subsets are presented in Figure 2. For the hg H tree, two data sets of different size (237 and 337 sequences) from different random sets of modern sequences were analysed to test the impact of the number of sequences and the subsampling process on the Bayesian calculation. No substantial difference was observed between the two replicates, as the mean rate estimate of each analysis is within the 95% highest posterior density (95% HPD) of its respective replicate. However, within the range of 95% HPD, the rates estimated from the larger data set appear systematically higher than the rates estimated from the smaller data set. For the whole AMH tree (Figure 2), four estimates are compared (see Table 2). No significant difference in estimated rates was observed between the different data sets. Moreover, the rate estimates have narrower 95% HPD than for calculations based on the hg H, and no particular data set shows consistently higher rates than the others.

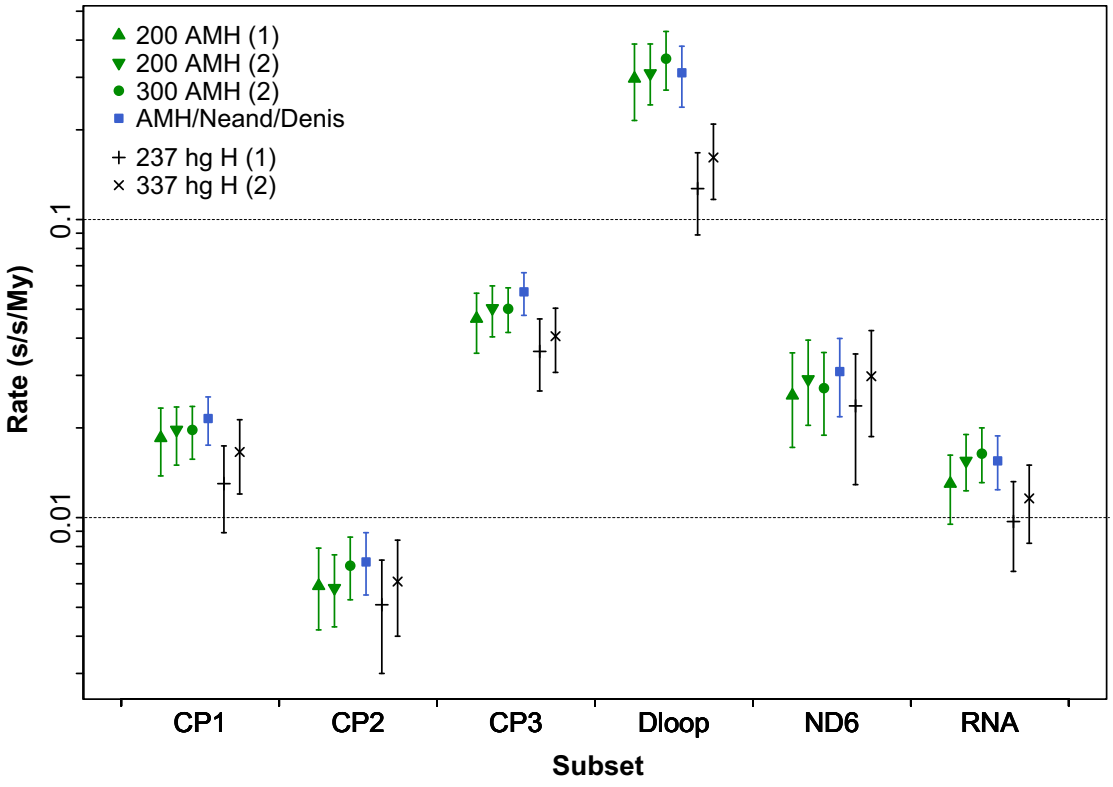


Figure 2. Comparison of rate estimates for the six mitochondrial subsets, calculated from six different datasets.

Interestingly, the comparison between rate estimates from hg H sequences only (black in Figure 2), or from sequences representing all AMH (coloured in Figure 2), reveals an unexpected difference. Although the time dependence of molecular rates would suggest that the rates calculated for the shorter timescale should be higher than for the deeper phylogeny, the rates inferred for hg H appear systematically lower than the ones inferred for the whole AMH tree. This difference is particularly noticeable for the 'fast evolving' subsets, such as

the D-loop and less so CP3 (Figure 2). When interpreted together with the previous observation that rates appear higher within hg H when estimated from more samples, it seems that the genetic signal (observed mutations) from the hg H subsamples, especially in the D-loop, does not accurately reflect the evolutionary speed of the entire AMH mitochondrial phylogeny. Even though more replicates will be required to obtain a finer pattern, the progression of rate estimates between a dataset of 237 samples and a dataset of 337 samples suggests that adding more sequences to the dataset also adds critical genetic signal for the inference, and that a 'correct' or representative rate value has not been reached. On the contrary, the number of sequences used to calculate the rate of the whole AMH phylogeny does not have any impact on the resulting value. That might be explained by the fact that when deep nodes are included in the phylogeny (such as the root of haplogroups L3, M, N and R), there is sufficient genetic change along the internal branches to obtain a reliable rate estimate, regardless of the size of the AMH dataset. As a consequence, this also means that the addition of further haplotypes at the tips does not provide crucial genetic signal.

The consistency of rate estimates between data sets, including or excluding Neandertal and Denisovan sequences, shows that the violation of the assumption of panmixia does not impact the results. However, for the inferences calibrated with the human-chimpanzee split (Supplementary table 1), we observe a striking difference between the rates estimated using only the coalescent model and the rates estimated with the *BEAST option. The mean rates estimated using *BEAST (e.g., a mean of 0.164 substitution per site per million year, s/s/My, [0.113 – 0.220 95% HPD], for the D-loop) are twice as high as the corresponding rate estimates based on the coalescent (e.g., a mean of 0.082 s/s/My, [0.064 – 0.101 95% HPD], for the D-loop). Moreover, the mean rates estimated from the human-chimpanzee fossil calibration, using the coalescent model, are lower than the corresponding estimates based on tip calibration (see Supplementary table 1), while the rate estimates obtained using *BEAST reach values similar to the rates calculated for the hg H data sets. When human and chimpanzee sequences are included we observe an underestimation of molecular rates. The use of the *BEAST option, with the implementation of a Yule model at the interspecific level, combined with the coalescent model to account for the intraspecific diversity, appears to efficiently correct for the observed bias. However, as expected from the time dependence of molecular rates, even using the *BEAST option, the estimated rates based on the human-chimpanzee fossil calibration are still lower than the rates inferred using tip calibrations.

Comparison of dates between data sets

The estimated dates for the major AMH haplogroups and the main H sub-haplogroups are reported Figure 3a and 3b. As expected from the estimated rates for each hg H dataset, the respective date estimates are consistent between both data sets, with slightly younger estimates for the larger dataset (hg H 337, Fig. 3a). For the entire AMH tree, dates are consistent between all four data sets, except for the root of the AMH clade (Fig. 3b). The data set including Neandertal and Denisovan samples leads to an older tMRCA estimate for AMH than the data sets including only AMH. Interestingly, a similar observation is made from the tMRCA of hg H data. When only samples from the nested phylogeny of hg H samples are considered, hg H tMRCA estimates appear younger than those based on sequences from all major AMH haplogroups. We observe that the rooting of a particular node with outgroup sequences (Neandertal/Denisovan sequences on one hand, and non-H sequences on the other hand) is clearly influencing its age estimate, i.e. making it older, while the dates in the nested phylogeny remain unaffected. This could be the result of a better distribution of the inferred mutation events towards the basal branches of the tree, by adding genetic diversity outside the group of interest.

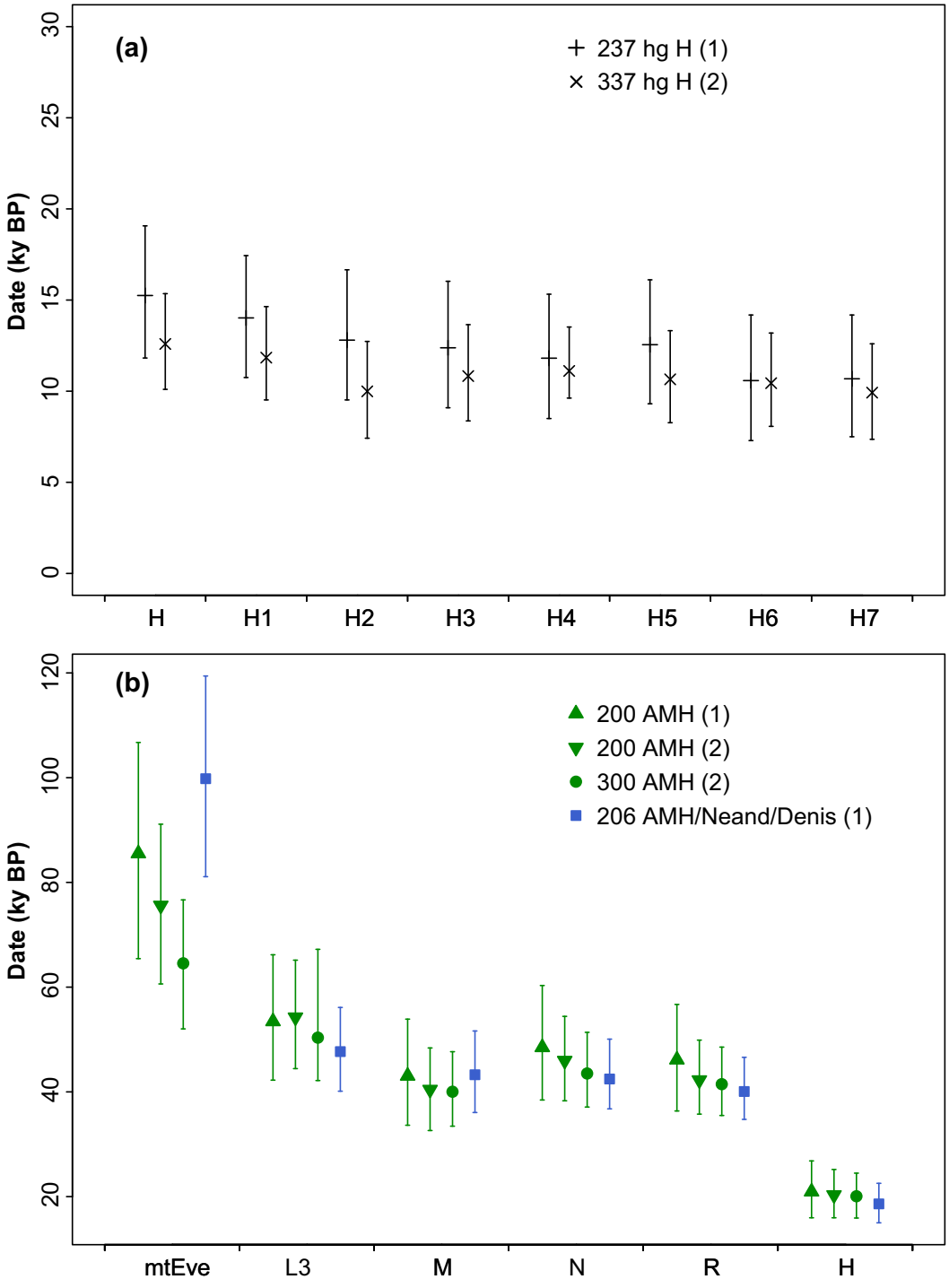


Figure 3. Comparison of coalescent age estimates for seven common sub-haplogroups of hg H **(a)**, and for the main macro-haplogroups of AMH **(b)**, calculated from different datasets.

Comparison to published calculations

When the partitioning scheme allows the comparison of our rate estimates with those of previous studies, the rates inferred using tip calibrations from ancient sequences are higher than the ones estimated using a human-chimpanzee calibration. For example, the widely cited rate for the human mitochondrial coding region from Mishmar and colleagues (2003) (1.26 $x10^{-8}$ s/s/y) is very similar to the rates calculated by Soares and colleagues (2009), which is also based on the human-chimpanzee split (0.88 $x10^{-8}$ s/s/y for coding positions 1 and 2, and 1.93 $x10^{-8}$ s/s/y for the third codon position). Estimates for the coding region from both studies are also lower than the rates estimated in the present study (Supplementary table 1). This trend is continued when rate estimates from D-loop and RNA portions are compared: 16.43, 22.96 and 1.55 $x10^{-8}$ s/s/y are inferred for different parts of the D-loop (Soares et al. 2009), whereas we estimate a rate of 29.72 (21.50-38.80) $x10^{-8}$ s/s/y for the entire D-loop. Similarly, rates of 0.82 and 0.69 $x10^{-8}$ s/s/y are estimated for rRNA and tRNA, respectively, by (Soares et al. 2009), whereas we estimate 1.30 (0.95-1.62) $x10^{-8}$ s/s/y for both of them combined.

While the rates estimated from ancient DNA calibrations in the present study are higher than commonly used ones from human-chimpanzee calibrations, they still are lower than mutation rates reported by pedigree studies (Howell et al. 2003; Santos et al. 2005; Madrigal et al. 2012). It is generally understood that rate estimates from pedigree studies include many mutations that will most likely not become fixed in future generations. Therefore, these 'fast' estimates represent the extreme end of a spectrum and can be considered overestimations when used as substitution rates.

In contrast, the rates obtained in our study are more similar to the rates previously inferred from internal calibrations (Endicott, Ho 2008). Here, rates were calculated for four subsets of the mitochondrial genome based on archaeological and biogeographical events like the entry of AMH in Australia and New Guinea. The rates estimates for the D-loop and CP3 are directly comparable with the present study and show strikingly similar values: 30.2 (22.3-37.3) and 5.09 (3.44-6.80) $x10^{-8}$ s/s/y, respectively, compared to 29.72 (21.50-38.80) and 4.65 (3.56-5.66) $x10^{-8}$ s/s/y for the present study (Supplementary table 1).

The comparisons of rate estimates across phylogenetic and pedigree studies are in accordance with the time dependence of molecular rates. The rates estimated from pedigree calculations are (far) higher than the equivalent rates estimated based deep fossil calibration using the human/chimpanzee split. Estimates from direct calibrations using ancient DNA or indirect calibration(s) based on archaeological evidences result in intermediate values for substitution rates. Rather than using estimates from either end of the spectrum, these direct

estimates are therefore likely to be most precise for the time periods that include the major events in the genetic history of AMH.

The dated phylogeny inferred from the AMH dataset 200 and dates estimated for the main haplogroups are presented in Figure 4 (phylogenies inferred from the other data sets are presented in Supplementary figures 1-5). The resulting topology of the main mitochondrial haplogroups is consistent with the one that is currently accepted (Behar et al. 2012; http://www.phylotree.org). African lineages, here represented by haplogroups L0, L1 and L2, are at the root of the tree. The haplogroup L3 represents the onset of the migration out of Africa, followed by the rapid diversification of the non-African human population into three main lineages: M, N and R.

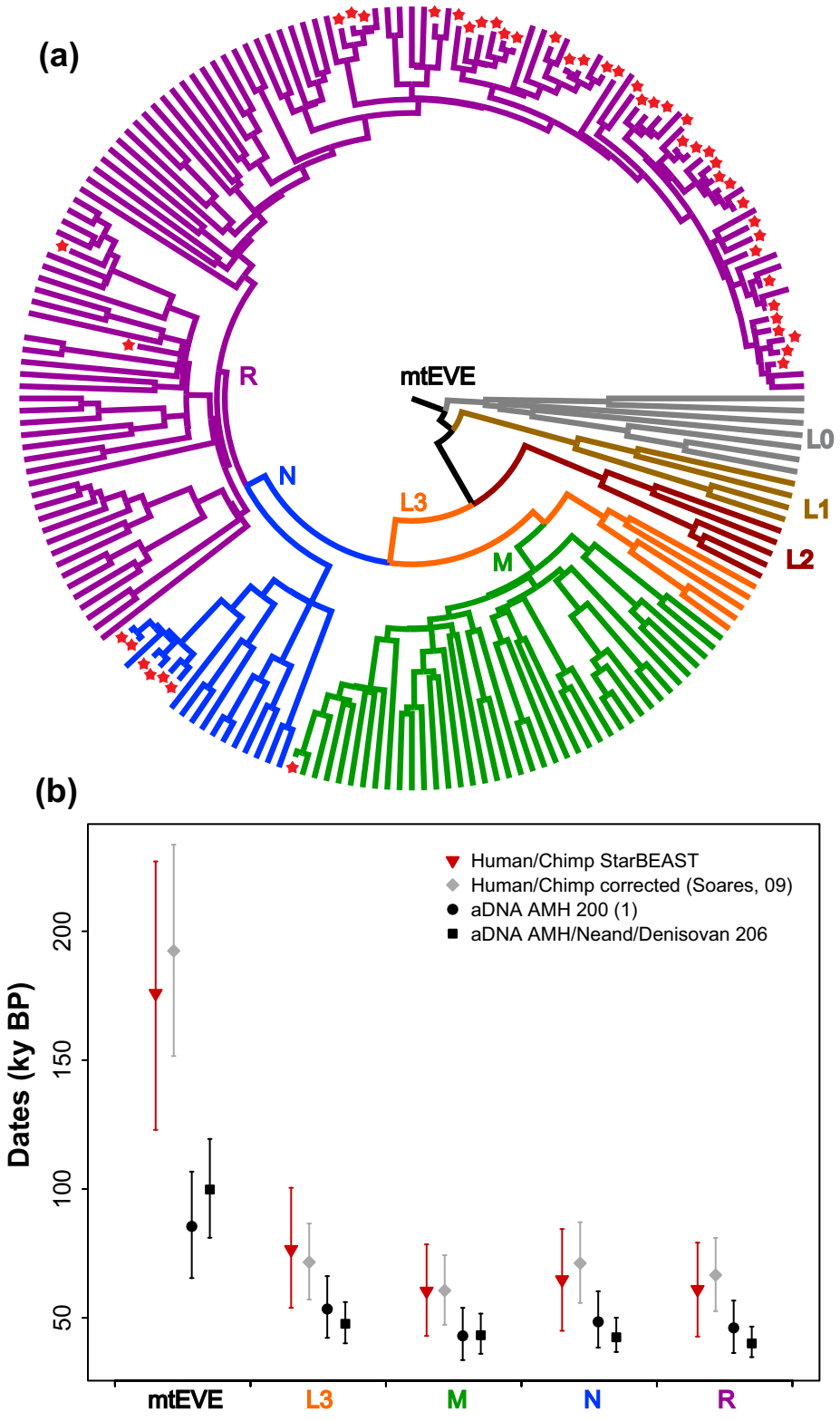


Figure 4. Example of an inferred AMH phylogeny and comparison of resulting coalescent age estimates. **(a).** Maximum clade credibility tree estimated using Bayesian analysis the data set AMH 200 (1). The red stars represent the ancient samples used to calibrate the tree. **(b).** Comparison of coalescent dates obtained from the tree presented above with two other calculations (*BEAST and AMH/Neandertal/Denisovan 206), and estimates from a previous study (Soares et al. 2009).

We replicated the Bayesian inference of the AMH tree using the human-chimpanzee split as calibration, in order to compare the effects of direct internal calibration (with aDNA) and external calibration (using fossil evidence) on the estimated timescale of human evolution. All dates estimated from tip calibrations are younger than the ones estimated from the human/chimpanzee calibration point. For example, the AMH tMRCA for the mitochondrial genome (labelled as 'mtEve' on the figures) is calculated to be at 99.8 ky BP (81.1 to 119.4) using all available tip calibration points (including Neandertal and Denisovan sequences), but is estimated at 175.9 ky BP (123 to 228.9) when fossil calibration is used. Using fossil calibration also places the date of the migration of AMH out of Africa (represented by the tMRCA of the L3 hg) around 75 ky BP (53.9 to 100.5), whereas this event is estimated to have happened only around 50 ky BP (42.2 to 66.2) when ancient DNA is used as calibration. Similarly, the diversification of hg L3 into the three main non-African haplogroups at the onset of the colonisation of other continents by AMH is calculated to have happened between 43 and 48 ky BP (33.6 to 60.3), in contrast to 60 to 65 ky BP (42.8 to 84.5) when estimated with fossil calibration.

Importantly, the present timescale based on the human-chimpanzee calibration is remarkably similar to estimates from a recently proposed human mitochondrial molecular clock (Soares et al. 2009). This clock takes the time dependence of molecular rates into consideration by correcting for purifying selection. However, this calculation still relies on the date estimate of the human-chimpanzee split, without taking its uncertainty into account. While this method corrects the rate for purifying selection, it ignores other potential causes of time-dependent rates (Ho et al. 2011a). In addition, the rate of each mitochondrial subset is calculated independently from a fixed topology, and the maximum likelihood method used does not take into account the rate heterogeneity among lineages. Furthermore, a similar timescale is obtained using a Bayesian approach without a correction for purifying selection (Figure 4). In contrast, one advantage of the Bayesian method is that it accounts for uncertainty in the topology, the demographic history and the six different molecular clocks (one for each subset) in the same Monte Carlo Markof Chain (MCMC) analysis. It can also take into account the uncertainty of the date of the human-chimpanzee split by applying a lognormal distribution to the calibration, and it allows the application of both a coalescent model for intraspecific level, and a Yule model, necessary to include chimpanzee taxa at the interspecific level. However, despite all methodological efforts to correct for the time dependence of molecular rates, there is still a strong discrepancy between date estimates based on external fossil calibration (Soares et al. 2009, and present *BEAST analysis) and those based on internal calibrations (aDNA).

The timescale inferred using ancient DNA in the present study is younger than previous inferences based on fossil calibration (Mishmar et al. 2003; Soares et al. 2009), but is in accordance with recent estimates based on internal calibrations (Endicott, Ho 2008). The resulting scenario for the recent AMH evolutionary history is a recent migration out of Africa (post-Toba), followed by a rapid diversification of non-African haplogroups and dispersal in South-East Asia and Europe (Endicott et al. 2009; Oppenheimer 2012).

Demographic reconstruction

The coalescent model used to estimate the evolutionary timescale of AMHs also allows the reconstruction of past effective population size. The Bayesian skyride plot presented in Figure 5 shows two periods of rapid population growth and a period of population decline (plots from all data sets are presented in Supplementary figures 1–5).The first expansion, between around 45 and 25 ky BP, directly follows the proposed exodus from Africa and marks the split of L3 into three major non-African lineages M, N and R. During this period, all continents except the Americas were colonised by lineages derived from the M macro-haplogroup through a southern route toward Asia, and lineages derived from the N macro-haplogroup through a northern route, including haplogroup R, which encompasses most Eurasian lineages. Between around 25 and 18 ky BP, the effective population size of AMH decreased, which is consistent with the Last Glacial Maximum (LGM) halting human expansions in higher latitudes. A second exponential growth of the human population is observed between the end of the LGM and today. The beginning of this expansion period corresponds to the lineage diversification after the LGM, represented here by seven subclades of hg H appearing strikingly simultaneous between 17 and 9ky ago. In its later phases this post-LGM diversification and exponential population growth coincides with the Neolithic transition (starting from ~12 ky BP in the Fertile Crescent in the Near East, and ~9 ky in China for example, Diamond, Bellwood 2003).

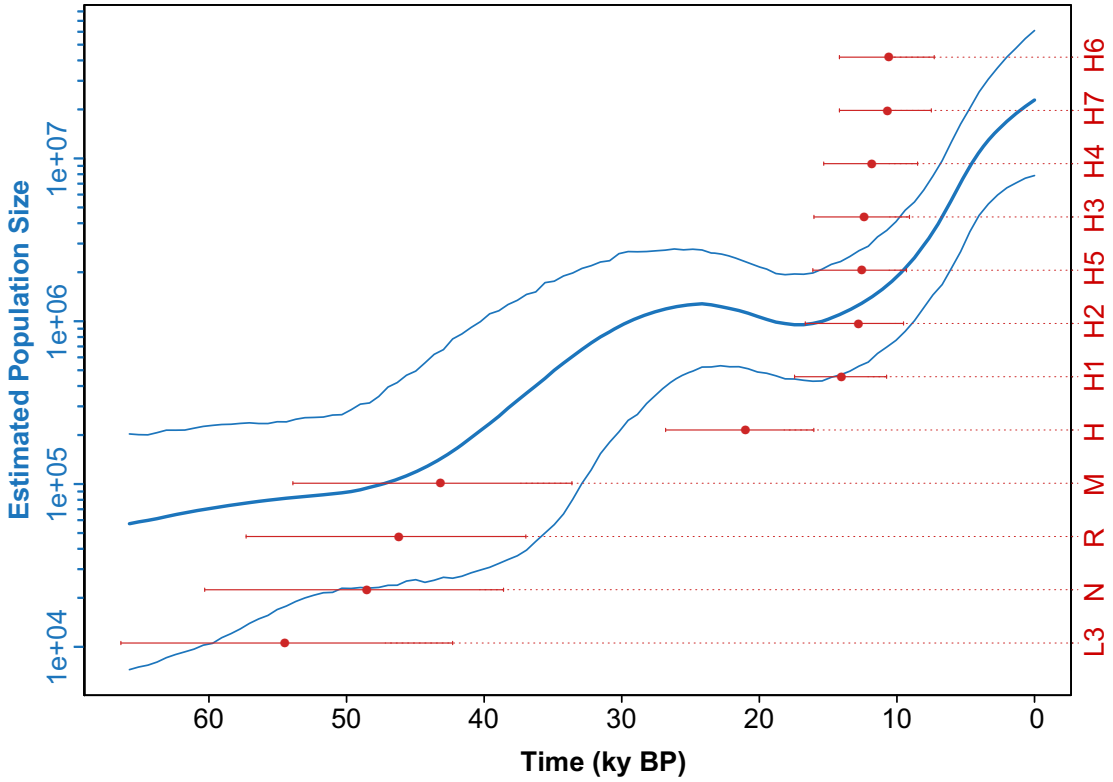


Figure 5. Bayesian skyride plot of effective population size through time (in blue), and 12 coalescent age estimates of the AMH phylogeny (in red). All estimates reported on this figure were calculated with the data set AMH 200 (1), except for the hg H coalescent dates, calculated with the data set hg H 237.

It should be noted that the sampling of human mtDNA haplogroups in our study is biased toward European lineages. Consequently, although the present study highlights the potential of current methods to infer human demographic history using ancient sequences as internal calibrations, further analyses with an even coverage of the human mtDNA phylogeny are desirable to refine the results and to focus on regional population and peopling events. The limiting factor will most likely be the availability of dated ancient sequences to obtain a more detailed coverage of AMH sub-haplogroup lineages, and to represent the populations of interest.

CONCLUSION

With the present study, we report the first dated phylogeny of modern humans internally calibrated using the date associated with ancient whole mitochondrial genomes. We demonstrate that although the 43 available ancient WMG only cover a small part of the AMH haplotype diversity, they contain enough temporal signal to calibrate the AMH phylogeny. Finally, we show that Bayesian phylogenetic methods offer a powerful framework to integrate the temporal signal from heterochronous sequences and to co-estimate the tree topology, the substitution rate of the different mitochondrial subsets and the ages of coalescent events in the same analysis.

Overall, we observe remarkably consistent estimates of rates and dates from different sampling approaches varying in size and composition. This confirms that the Bayesian approach used in this study is capable of inferring accurate estimates from a subsampling of ~200 random sequences from all haplogroups. However, the observed pattern of slightly higher rates and younger dates calculated for hg H when more samples are used does suggest a sampling effect at the haplogroup level that might be further explored in future investigations.

References

Atkinson, QD, RD Gray, AJ Drummond. 2008. mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. Molecular Biology and Evolution 25:468-474.

Bandelt, HJ, QP Kong, YG Yao, M Richards, V Macaulay. 2006. Mitochondrial DNA and the Evolution of Homo sapiens. Berlin: Springer-Verlag.

Behar, DM, M van Oven, S Rosset, M Metspalu, EL Loogvali, NM Silva, T Kivisild, A Torroni, R Villems. 2012. A "copernican" reassessment of the human mitochondrial DNA tree from its root. American Journal of Human Genetics 90:675-684.

Briggs, AW, JM Good, RE Green, et al. 2009. Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science 325:318-321.

Diamond, J, P Bellwood. 2003. Farmers and their languages: the first expansions. Science 300:597-603.

Drummond, AJ, OG Pybus, A Rambaut, R Forsberg, AG Rodrigo. 2003. Measurably evolving populations. Trends in ecology & evolution (Personal edition) 18:481-488.

Drummond, AJ, A Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC evolutionary biology 7:214.

Endicott, P, SYW Ho. 2008. A Bayesian evaluation of human mitochondrial substitution rates. American Journal of Human Genetics 82:895-902.

Endicott, P, SYW Ho, M Metspalu, C Stringer. 2009. Evaluating the mitochondrial timescale of human evolution. Trends in ecology & evolution 24:515-521.

Ermini, L, C Olivieri, E Rizzi, et al. 2008. Complete mitochondrial genome sequence of the Tyrolean Iceman. Current biology : CB 18:1687-1693.

Gilbert, MT, T Kivisild, B Gronnow, et al. 2008. Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland. Science 320:1787-1789.

Gouy, M, S Guindon, O Gascuel. 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular Biology and Evolution 27:221-224.

Green, RE, AS Malaspinas, J Krause, et al. 2008. A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. Cell 134:416-426.

Henn, BM, CR Gignoux, MW Feldman, JL Mountain. 2009. Characterizing the time dependency of human mitochondrial DNA mutation rate estimates. Molecular Biology and Evolution 26:217-230.

Ho, SYW, R Lanfear, L Bromham, MJ Phillips, J Soubrier, AG Rodrigo, A Cooper. 2011a. Time-dependent rates of molecular evolution. Molecular ecology 20:3087-3101.

Ho, SYW, R Lanfear, MJ Phillips, I Barnes, JA Thomas, SO Kolokotronis, B Shapiro. 2011b. Bayesian estimation of substitution rates from ancient DNA sequences with low information content. Systematic biology 60:366-375.

Ho, SYW, MJ Phillips, A Cooper, AJ Drummond. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Molecular Biology and Evolution 22:1561-1568.

Howell, N, CB Smejkal, DA Mackey, PF Chinnery, DM Turnbull, C Herrnstadt. 2003. The pedigree rate of sequence divergence in the human mitochondrial genome: There is a difference between phylogenetic and pedigree rates. American Journal of Human Genetics 72:659-670.

Krause, J, AW Briggs, M Kircher, T Maricic, N Zwyns, A Derevianko, S Paabo. 2010a. A complete mtDNA genome of an early modern human from Kostenki, Russia. Current biology : CB 20:231-236.

Krause, J, Q Fu, JM Good, B Viola, MV Shunkov, AP Derevianko, S Paabo. 2010b. The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. Nature 464:894-897.

Lanfear, R, B Calcott, SYW Ho, S Guindon. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution.

Madrigal, L, LC Posthumously, M Melendez-Obando, R Villegas-Palma, R Barrantes, H Raventos, R Pereira, D Luiselli, D Pettener, G Barbujani. 2012. High mitochondrial mutation rates estimated from deep-rooting costa rican pedigrees. American journal of physical anthropology.

Minin, VN, EW Bloomquist, MA Suchard. 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Molecular Biology and Evolution 25:1459-1471.

Mishmar, D, E Ruiz-Pesini, P Golik, et al. 2003. Natural selection shaped regional mtDNA variation in humans. Proceedings of the National Academy of Sciences of the United States of America 100:171-176.

Oppenheimer, S. 2012. Out-of-Africa, the peopling of continents and islands: tracing uniparental gene trees across the map. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 367:770-784.

Palanichamy, MG, CL Zhang, B Mitra, B Malyarchuk, M Derenko, TK Chaudhuri, YP Zhang. 2010. Mitochondrial haplogroup N1a phylogeography, with implication to the origin of European farmers. BMC evolutionary biology 10:304.

Santos, C, R Montiel, B Sierra, C Bettencourt, E Fernandez, L Alvarez, M Lima, A Abade, MP Aluja. 2005. Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: a model using families from the Azores Islands (Portugal). Molecular Biology and Evolution 22:1490-1505.

Soares, P, L Ermini, N Thomson, M Mormina, T Rito, A Rohl, A Salas, S Oppenheimer, V Macaulay, MB Richards. 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. American Journal of Human Genetics 84:740-759.

Stewart, JR, CB Stringer. 2012. Human evolution out of Africa: the role of refugia and climate change. Science 335:1317-1321.

Stoneking, M, ST Sherry, AJ Redd, L Vigilant. 1992. New Approaches to Dating Suggest a Recent Age for the Human mtDNA Ancestor. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 337:167-175.

Subramanian, S, DR Denver, CD Millar, T Heupink, A Aschrafi, SD Emslie, C Baroni, DM Lambert. 2009. High mitogenomic evolutionary rates and time dependency. Trends in Genetics 25:482-486.

van Oven, M, M Kayser. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Human mutation 30:E386-394.

Supplementary material

Supplementary table 1. Detail of the molecular rate estimates for the six mitochondrial subsets, calculated from 8 different datasets. Rate values are given in $x10^{-8}$ s/s/y.

	AMH 200 (1)			AMH 200 (2)			AMH 300			AMH/Neand/Denis 206		
CP1	**1.85**	1.38	2.33	**1.97**	1.50	2.35	**1.97**	1.57	2.36	**2.15**	1.75	2.54
CP2	**0.59**	0.42	0.79	**0.58**	0.43	0.75	**0.69**	0.53	0.86	**0.71**	0.55	0.89
CP3	**4.65**	3.56	5.66	**5.04**	4.04	5.99	**5.02**	4.18	5.90	**5.72**	4.77	6.63
Dloop	**29.72**	21.50	38.80	**30.94**	24.25	38.82	**34.67**	27.18	42.76	**31.07**	23.82	38.15
ND6	**2.57**	1.72	3.57	**2.91**	2.04	3.94	**2.72**	1.89	3.58	**3.09**	2.18	3.99
RNA	**1.30**	0.95	1.62	**1.55**	1.23	1.90	**1.64**	1.31	2.00	**1.55**	1.24	1.88

	AMH/chimp 159 -			AMH/chimp 159 -		
CP1	**0.61**	0.49	0.73	**1.15**	0.80	1.54
CP2	**0.21**	0.17	0.26	**0.39**	0.26	0.53
CP3	**1.72**	1.42	2.05	**3.32**	2.43	4.39
Dloop	**8.22**	6.44	10.12	**16.41**	11.34	21.98
ND6	**1.02**	0.71	1.34	**1.82**	1.24	2.51
RNA	**0.45**	0.36	0.56	**0.86**	0.62	1.11

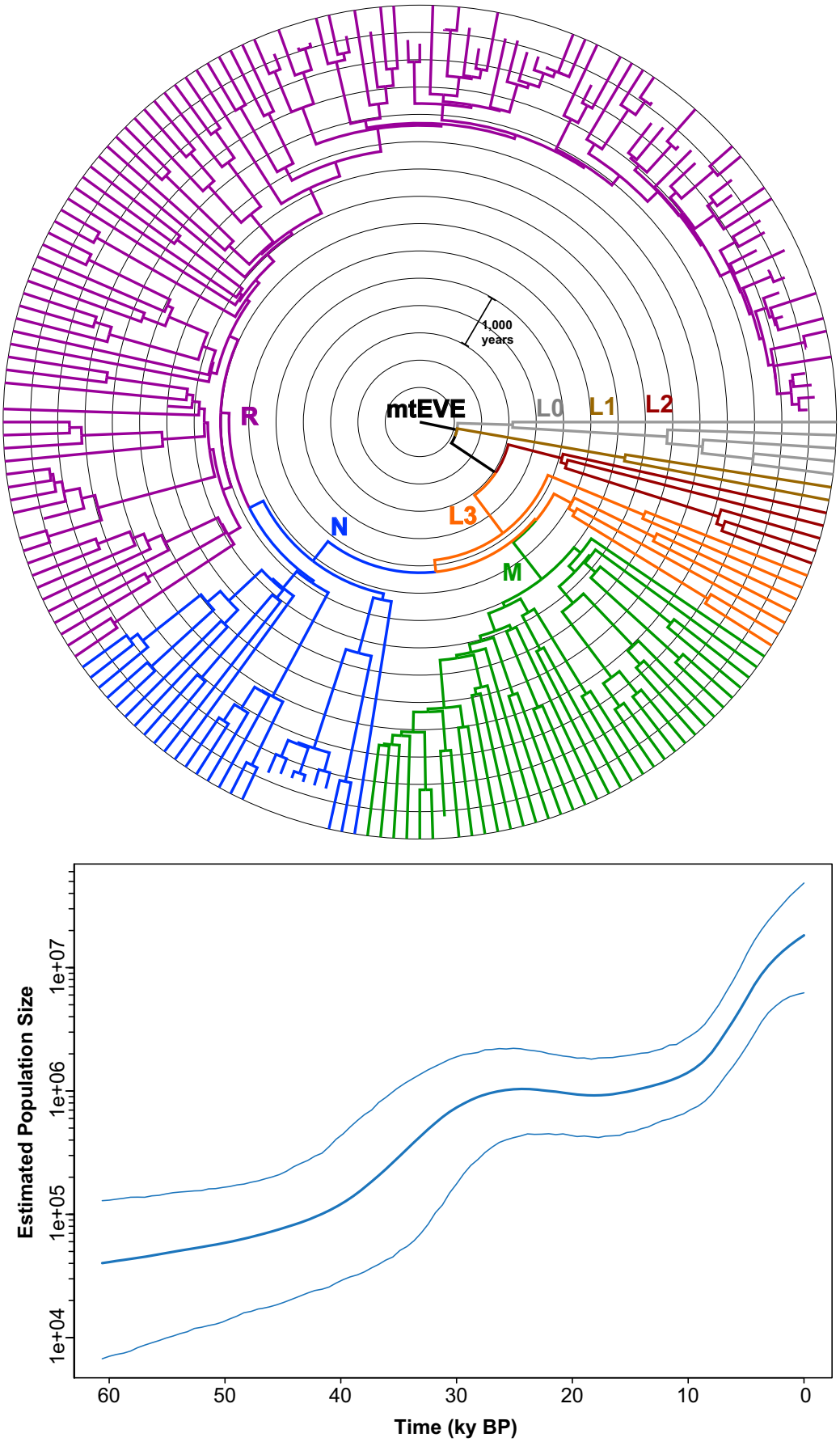
	hg H 237			hg H 337		
CP1	**1.30**	0.89	1.74	**1.66**	1.20	2.13
CP2	**0.51**	0.30	0.72	**0.61**	0.40	0.84
CP3	**3.61**	2.66	4.64	**4.06**	3.07	5.04
Dloop	**12.70**	8.88	16.75	**16.15**	11.68	20.90
ND6	**2.37**	1.29	3.54	**2.98**	1.87	4.24
RNA	**0.97**	0.66	1.32	**1.16**	0.82	1.50

Supplementary table 2. Detail of the coalescent age estimates for the main AMH haphogroups, and hg H sub-haplogroups, infered from 8 different datasets. Dates are in ky.

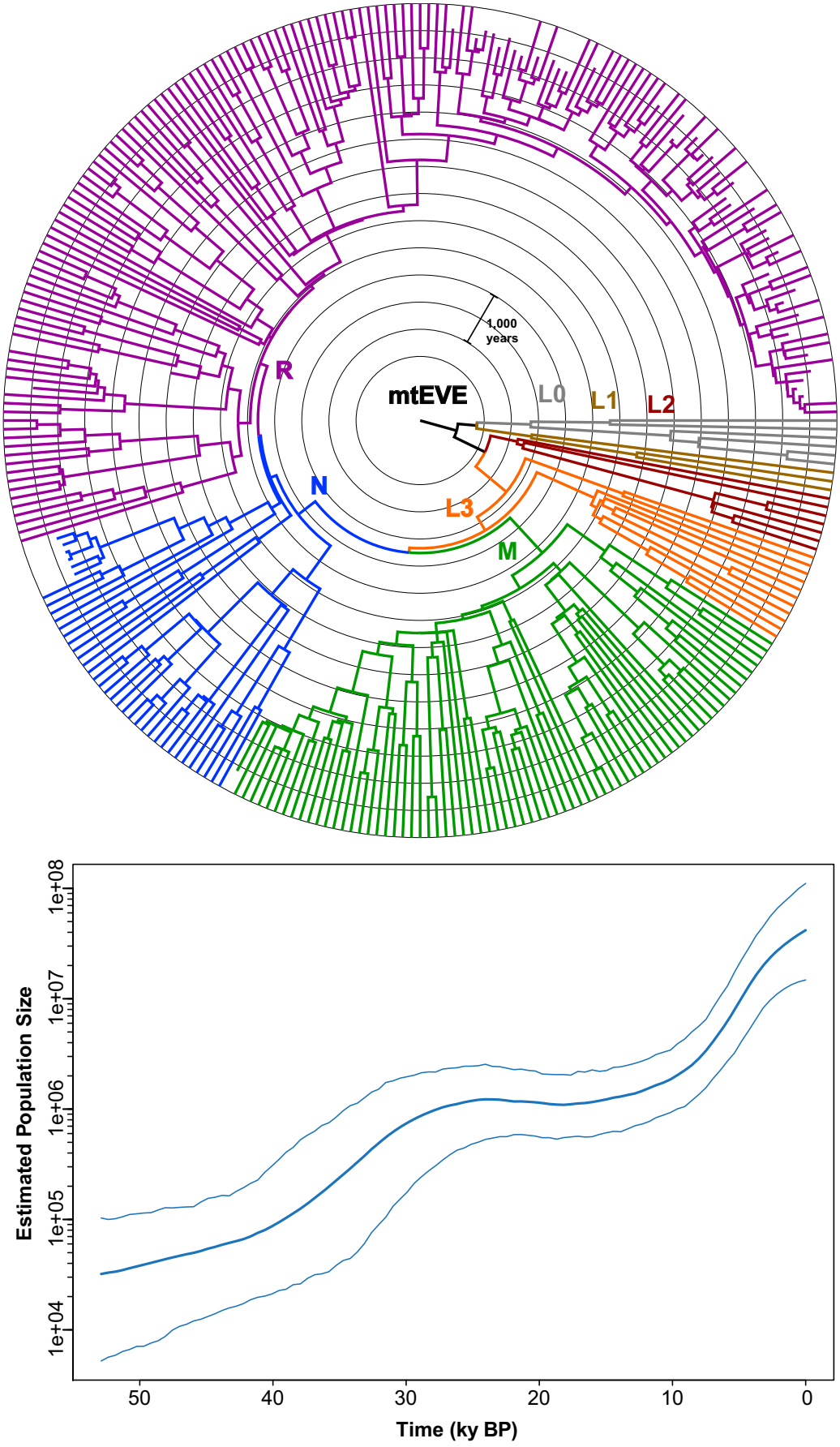
	AMH 200 (1)			AMH 200 (2)			AMH 300			AMH/Neand/Denis 206		
mtEve	**85.52**	65.44	106.72	**75.59**	60.61	91.11	**64.57**	52.03	76.66	**99.80**	81.09	119.41
L3	**53.45**	42.24	66.18	**54.25**	44.44	65.15	**50.37**	42.16	67.22	**47.69**	40.13	56.12
M	**43.05**	33.63	53.89	**40.44**	32.62	48.39	**40.05**	33.45	47.67	**43.27**	36.06	51.63
N	**48.50**	38.45	60.31	**45.97**	38.30	54.42	**43.53**	37.10	51.37	**42.46**	36.76	50.07
R	**46.14**	36.35	56.70	**42.25**	35.75	49.88	**41.49**	35.49	48.55	**40.07**	34.75	46.60
H	**20.94**	15.96	26.82	**20.31**	15.96	25.18	**20.09**	15.91	24.48	**18.61**	15.00	22.54

	AMH/chimp 159 -			AMH/chimp 159 -		
mtEve	**334.02**	263.32	405.50	**175.91**	122.95	227.08
L3	**161.75**	128.02	195.34	**76.52**	53.90	100.47
M	**125.75**	100.89	153.20	**60.36**	43.00	78.53
N	**142.34**	114.87	171.78	**64.82**	44.99	84.47
R	**134.18**	107.61	160.60	**61.02**	42.75	79.17
H	**58.46**	43.82	75.51	**29.49**	18.85	40.41

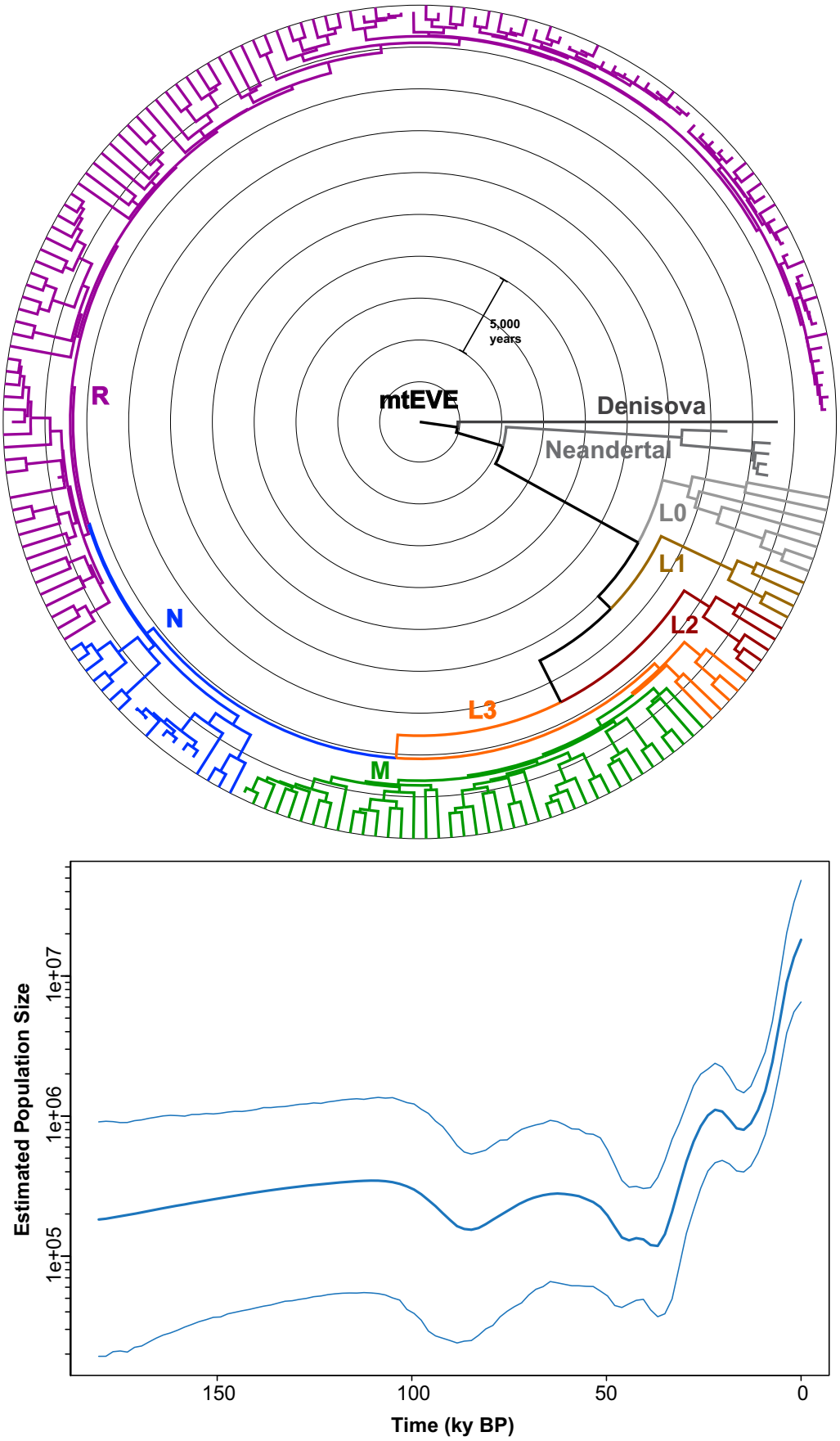
	hg H 237			hg H 337		
H	**15.25**	11.82	19.07	**12.59**	10.10	15.35
H1	**14.02**	10.75	17.44	**11.84**	9.52	14.64
H2	**12.80**	9.52	16.66	**9.99**	7.42	12.73
H3	**12.38**	9.09	16.03	**10.83**	8.37	13.65
H4	**11.81**	8.50	15.32	**11.11**	9.62	13.52
H5	**12.55**	9.31	16.11	**10.65**	8.27	13.32
H6	**10.58**	7.29	14.18	**10.44**	8.07	13.19
H7	**10.68**	7.50	14.18	**9.92**	7.36	12.60



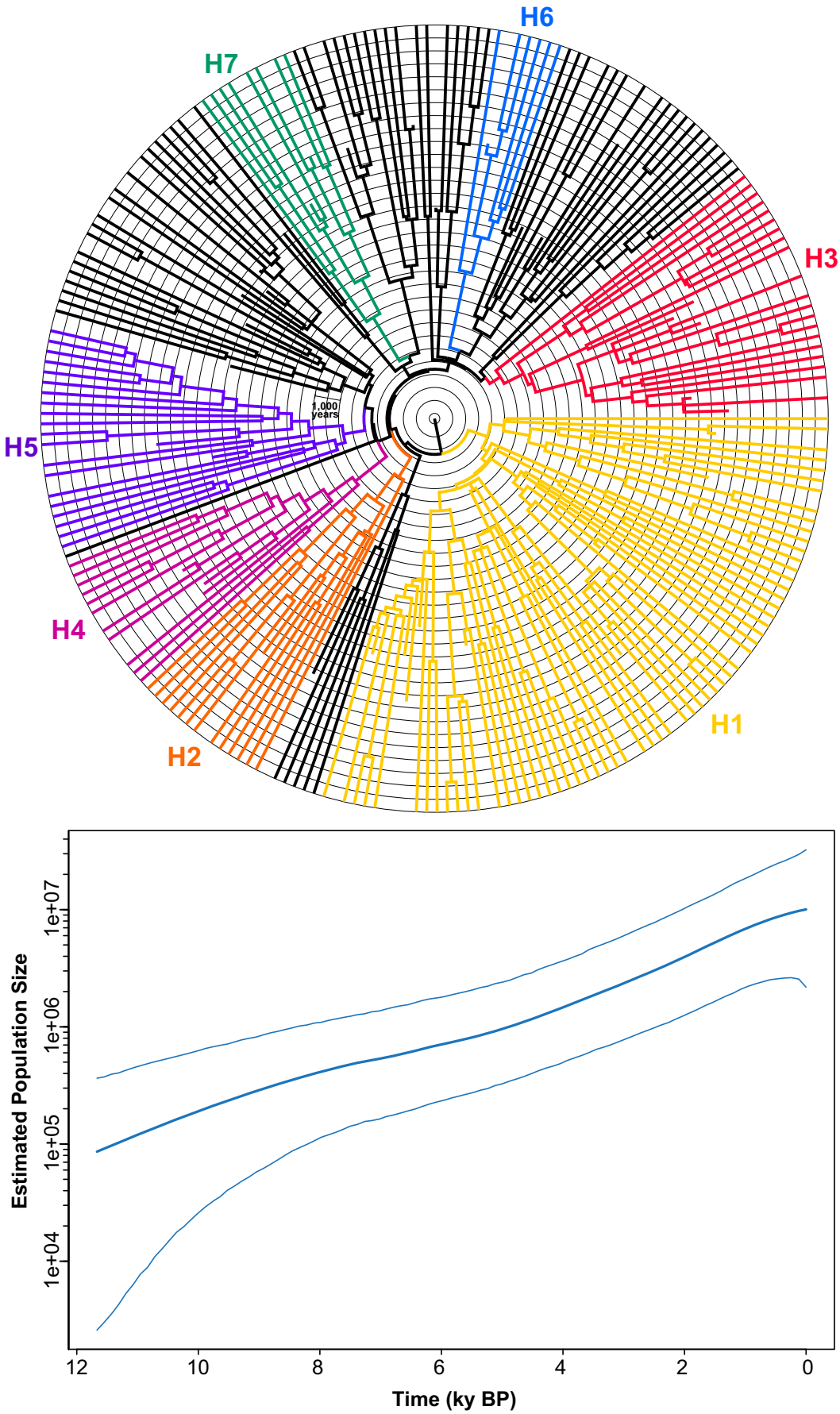
Supplementary figure 1. Maximum clade credibility tree estimated using Bayesian analysis of 200 AMH (2) WMG and the corresponding Bayesian skyride plot of effective population size through time.



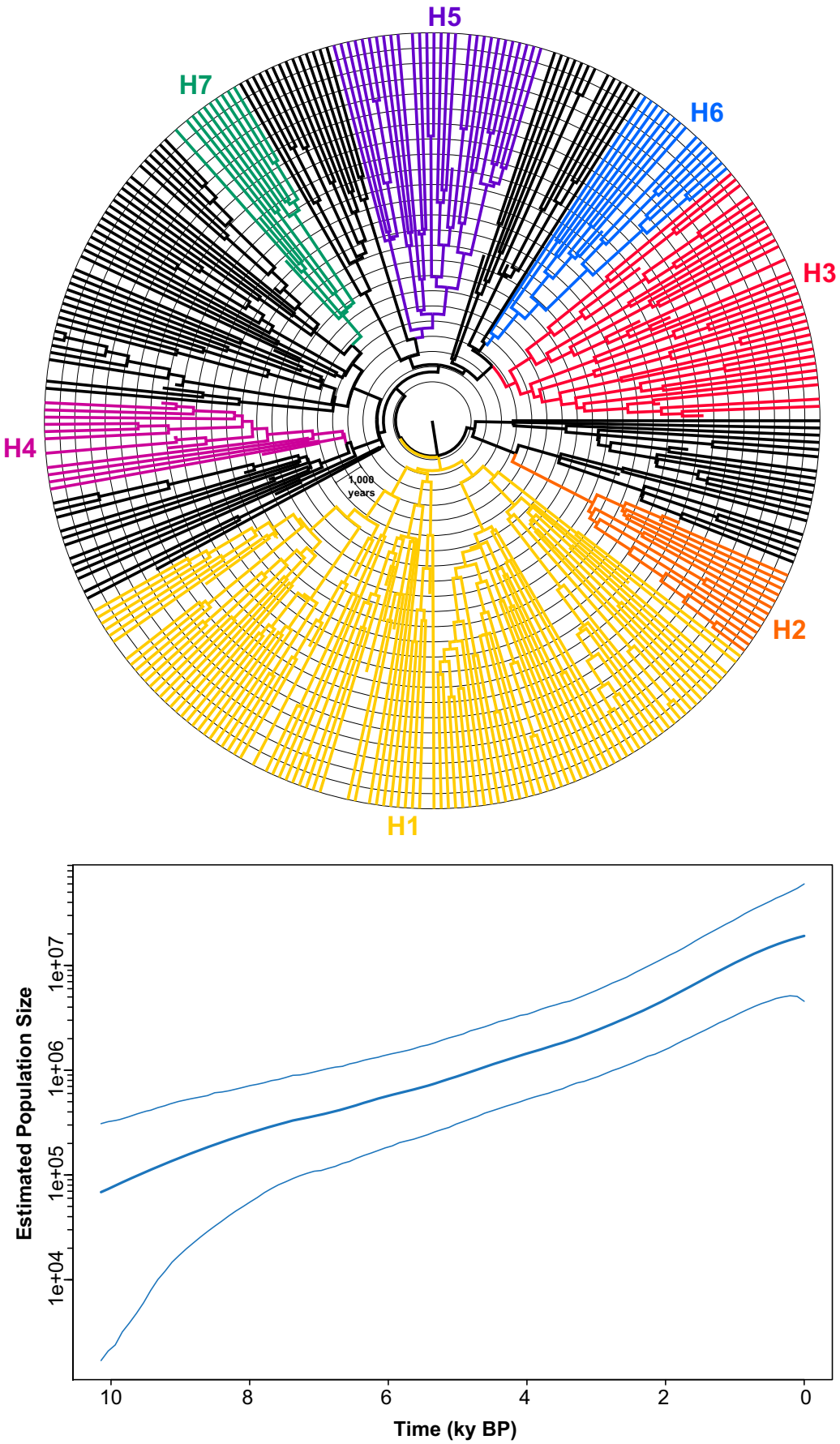
Supplementary figure 2. Maximum clade credibility tree estimated using Bayesian analysis of 300 AMH WMG and the corresponding Bayesian skyride plot of effective population size through time.



Supplementary figure 3. Maximum clade credibility tree estimated using Bayesian analysis of 206 AMH, Neandertal and Denisovan WMG and the corresponding Bayesian skyride plot of effective population size through time.



Supplementary figure 4. Maximum clade credibility tree estimated using Bayesian analysis of 237 hgH WMG and the corresponding Bayesian skyride plot of effective population size through time.



Supplementary figure 5. Maximum clade credibility tree estimated using Bayesian analysis of 337 hgH WMG and the corresponding Bayesian skyride plot of effective population size through time.

Chapter 7

Conclusion and future directions

Summary

This study has investigated the consequences of using ancient DNA to calibrate the molecular clock. Due to the temporal dependence of molecular rates, emphasis has been placed on the important role of such short term calibration points in calculating an accurate timescale of recent evolutionary history (e.g., population genetic studies, Quaternary events).

Most strikingly, Chapter 2 (with hyena) and Chapter 6 (with human) both contain direct comparisons of intra-specific evolutionary timescales calculated using either relatively old fossil (external node) calibrations or relatively recent tip (aDNA) calibrations. In both cases, the results confirmed previous observations that high rates are apparent when dated aDNA sequences are used as calibration points (Ho, Kolokotronis, Allaby 2007; Ho et al. 2011). Furthermore the differences in rate and date estimates from calibrations of different types and ages demonstrate the extent to which the temporal dependence of molecular rate estimates can lead to incompatible evolutionary scenarios. These observations question the reliability of any shallow molecular timescale inferred from old calibration point(s). Ancient DNA has proven to be a crucial source of genetic and temporal information to accurately infer recent evolutionary history, as it fills a temporal gap in calibration points, between fossil evidence (for old events) and genealogy studies (for much more recent events).

The dates associated with ancient sequences can also be used directly, to obtain chronological reconstructions of past population distribution, as shown in Chapter 4, with the study of Steppe bison and BisonX in Europe. This particular study shows that such chronological reconstructions can also assist with phylogenetic inference by refining the development of plausible scenarios of population replacement through time. Furthermore, the unprecedented precision of the timescales obtained using aDNA allows the correlation of past genetic changes with environmental events (e.g., climate change), as presented in both the European bovid megafauna study (Chapter 4) and the human mitochondrial study (Chapter 6, Part 2).

In the first part of Chapter 5, a review of the temporal dependence of molecular rate estimates is presented (Part 1). The main outcome of this work was a comprehensive presentation of the factors involved, and their potential impacts on the reconstruction of evolutionary timescales. The study presented the complexity of the subject and the multifaceted causes, and highlighted that the precise role of among-site rate heterogeneity and saturation was yet to be explored. This issue was therefore studied using mathematical models and simulated data in the second part of Chapter 5. Among other conclusions, this analysis showed the importance of data partitioning and/or complex evolutionary models, such as mixture models, to obtain accurate temporal results from the molecular clock. This result was exploited in Chapter 6 (Part 2), where the program PartitionFinder was used to empirically

define the optimal partitioning scheme for the human whole mitochondrial genome alignments of the study.

Based on that partitioning scheme, a re-calibration of the human mitochondrial tree is presented in Chapter 6, using the dates associated with 46 ancient whole mitochondrial genomes. This study shows the capacity of ancient mitogenomes to calibrate the entire human phylogeny when analysed with the latest Bayesian phylogenetic methods. But it also shows some potential limitations, with unexpectedly low rates inferred for hg H compared to the entire human tree, or the impact of outgroup sequences on the tMRCA estimates. The human timescale study also demonstrated the consequences of violating the assumptions of the coalescent model, such as when sequences from different species are used in the same analysis. This was illustrated through the use of the Human/Chimpanzee divergence as a calibration point, where the resulting rates and dates where in conflict if a (single-lineage) coalescent model was used for the entire phylogeny, versus the *BEAST multispecies coalescent approach. The *BEAST option acts as a 'multi-coalescent' model, implementing both the coalescent for intra-specific genetic diversity, and the Yule or birth-death model at the inter-specific level. The erroneous assumption of panmixia between humans and chimpanzee, required when a single coalescent was used, led to strong rate underestimates and consequently date overestimates for the main events in the evolution of anatomically modern humans. The use of the *BEAST option corrected for that bias, and led to a human evolutionary timescale similar to a key publication from 2009 that used corrections for purifying selection (Soares et al. 2009). However, molecular rate estimates based on tip calibration (using aDNA sequences) are persistently higher than estimates based on the human/chimpanzee split, regardless of the corrections applied (purifying selection, coalescent model, partitioning, relaxed clock, *etc*). This observation provides further evidence for the temporal dependence of molecular rate estimates, and disqualifies the distant fossil record as a reliable calibration method for inferring the timescale of recent human evolution using current methods.

On the other hand, the fact that molecular rate estimates are not transposable through time also means that phylogenetic inferences calibrated at the tip will not be suitable for estimating deep (inter-specific) divergence times. For example, this situation is observed in the second part of Chapter 5 (simulation study), where the time dependent bias analysed is of higher magnitude for nodes distant from the calibration point, whether for deep calibration or for shallow calibration. For that reason, situations for which the dates of interest are of intermediate age might ideally require the use of both fossil and tip calibration points, accommodated by a relaxed clock. But no methodological framework has been designed specifically for that situation, as discussed in the last part of the conclusion.

In Chapter 3, the typing of nuclear SNPs from ancient samples using existing commercial microarrays was investigated using bison samples, and the resulting paleogenomic data analysed in a phylogenetic framework. Along with the demonstration of the feasibility of such analysis, this study raised several technical and methodological issues. For example, heterozygote characters produced by SNP calling methods were shown to contain crucial phylogenetic signals, although in ancient samples a significant artifactual bias exists. As a result, an appropriate substitution model should be used to incorporate the heterozygote calls into phylogenetic analyses, whether using ancient or modern taxa. Most importantly, the accessibility of genome-wide nuclear information proved to be crucial in accurately reconstructing the evolutionary history of bison species, with the discovery of a previously unknown deep genetic divergence between the two surviving populations of American bison. The SNP data thus confirmed important discrepancies between mitochondrial and nuclear phylogenetic signals, as summarised in Figure C1. It is important to note that most ancient DNA studies remain focussed on mitochondrial targets (for understandable logistic reasons), and therefore this study has an important role in reiterating that phylogenetic reconstructions based on a single locus are not fully representative of the organisms' evolutionary history.

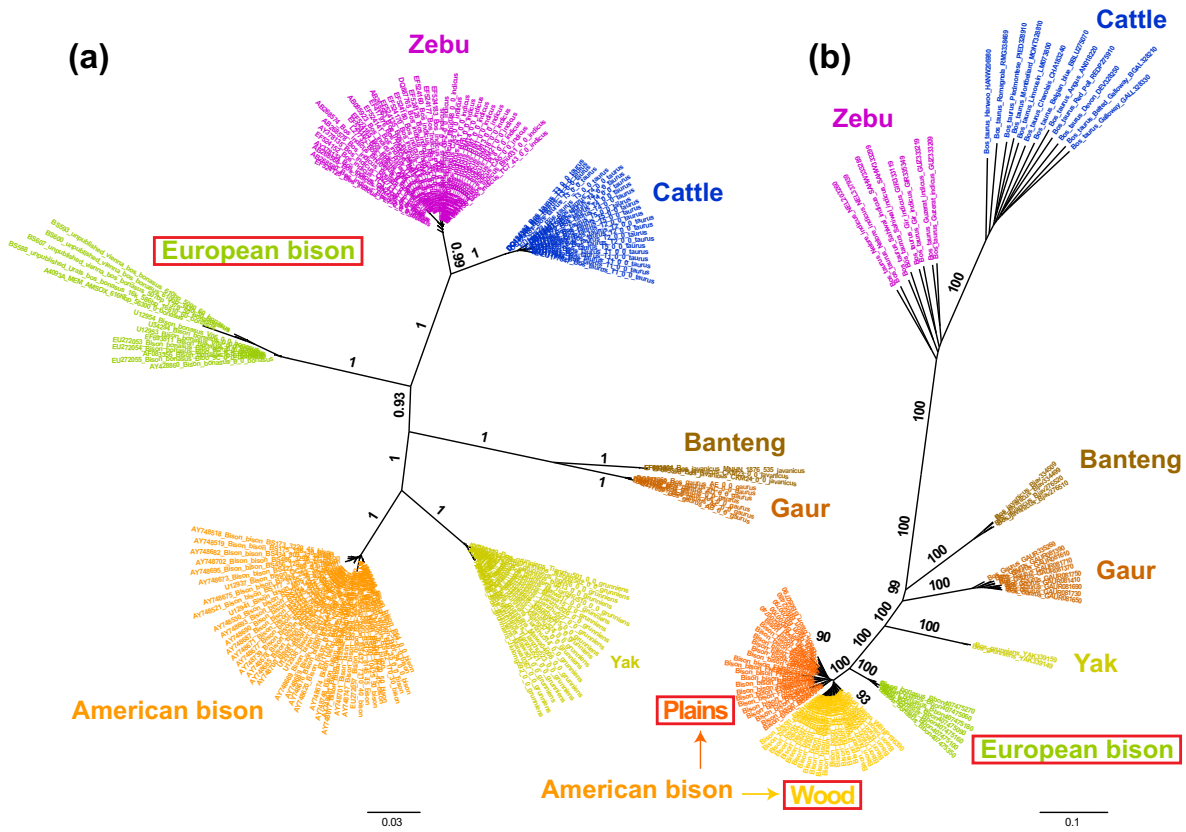


Figure C1. Comparison of mitochondrial (a) and nuclear (b) bovid phylogenies. The red boxes mark the main topological discrepancies mentioned in the text. **(a).** Maximum likelihood phylogeny of 152 bovid mitochondrial D-loop (627 bp) performed with PhyML, TN93 + G6 substitution model, and best of NNI + SPR searches. The branch support values are aLRT statistics. **(b).** Maximum likelihood phylogeny of 92 bovid nuclear SNP libraries (BovineSNP50) performed with RaxML, multistate GTR-like + G6 substitution model, 20 ML searches and 100 rapid Bootstrap (see Chapter 3 for methodological details).

In the light of the research presented in this thesis, it appears essential that molecular clock inferences follow certain protocols to avoid known biases and obtain reliable date estimates. Among other points, such protocols would include:

- The use of appropriate calibration(s). Due to the time dependence of molecular rates, calibration points should be chosen within a similar timeframe to the targeted divergence estimates (e.g., the fossil record is unlikely to provide reliable temporal information for population genetic studies).
- An objective selection of substitution models based on existing statistical tools, such as the comparison of Bayesian Information Criteria, as available in ModelGenerator (Keane et al. 2006) or jModelTest (Posada 2008).
- The selection of a partitioning scheme based on statistical comparisons, as available in PartitionFinder (Lanfear et al. 2012). The partitioning of sequence alignments is important to take into account the heterogeneity of evolutionary processes behind each character (like substitution model and rate). The use of mixture models can also help modelling such heterogeneity. This point is of particular importance considering the

increasing number of studies using phylogenomic data, including ancient DNA studies.

- Checking the robustness of the inferred topology, using statistical branch support calculations (like the bootstrap or posterior probabilities), but also by verifying the consistency of results between different phylogenetic reconstruction methods (like Bayesian and Maximum likelihood).
- The use of appropriate demographic models and priors when performing population level studies. In BEAST, this can be done using skyline/skyride models to account for potential demographic changes through time. For inter-specific studies, Yule or Birth-death priors can be used, while the *BEAST option allows the combination of inter and intra-specific genetic diversity in the same analysis (Drummond et al. 2012).
- The use of appropriate clock priors. In addition to the strict molecular clock, different degrees of relaxation of the clock are possible (correlated/uncorrelated relaxed clocks, random local clock).
- The use of several genetic loci. In addition to the simple increase of datapoints, the use of different genetic markers will help overcome topological and temporal biases due to incomplete lineage sorting, as well as other issues like population bottlenecks. The *BEAST option allows the use of multiple loci in the same MCMC analysis (Heled, Drummond 2010).
- Using a date randomization test to determine that dates associated with ancient DNA sequences provide enough temporal information to calibrate the phylogeny.
- Including calibration uncertainties by setting a distribution prior for calibration points, instead of a precise date. This can be performed either for fossil or tip calibrations when the precise date of the ancient sample is not known.

Future directions

As for most phylogenetic studies, the use of larger numbers of samples and/or longer sequences would help refine the results presented in the different chapters of this thesis. As an example, the results presented in Chapter 6 show that the human/chimpanzee divergence date does not provide a reliable calibration to estimate the recent human evolutionary history. On the other hand, ancient human whole mitochondrial genomes provide an unique opportunity to internally calibrate the human phylogenetic tree. However, the ancient mitogenomes available only represent a small proportion of the total number of human haplogroups, and better coverage of the human diversity with dated sequences would undoubtedly considerably refine the timescale and demographic inferences.

For both the hyena and bison mitochondrial studies the ancient DNA analyses are based on short control region fragments, and therefore the sequencing of whole mitochondrial genomes could help to confirm and refine the results obtained. But most importantly, the bison nuclear SNP study (Chapter 3) confirmed that inferences based exclusively on mitochondrial markers might not accurately represent the species' evolutionary history. Therefore, nuclear SNP chips or genomic studies certainly hold stronger potential to unravel the species histories and obtain accurate estimates of evolutionary timescales. However, the bison study has also shown that the genotyping of ancient samples using commercially available nuclear SNP chips was relatively inefficient, and further developments are required to generate reliable ancient libraries. New methods are currently under development to improve the quality of SNP genotyping of ancient individuals by enzymatic pretreatment of the ancient DNA extracts. One possibility resides in the removal of uracil residues, which are the principal artefacts produced by molecular degradation through time. Also, endogenous aDNA can be enriched in the extract, either by targeting prokaryotic DNA with restriction enzymes, and thus reducing the risk of bacterial contamination, or by target-enrichment of the SNP regions using RNA probes.

Several genomic extracts of Steppe bison individuals (including the ones used in the study presented in Chapter 3) are being processed and genotyped using these methods in order to replicate the results obtained in Chapter 3, and confirm the evolutionary scenario proposed for the origin of American bison. Although further sampling and additional replicates are likely to help refine our understanding of bison evolution, our capacity to infer accurate time estimates is still dependent on the age of the calibration point(s) used, as shown throughout this thesis. For bovid evolution, one fossil calibration is available at the inter-specific level between Yak and American bison, around 2.5 My (see Chapter 3). At the intra-specific level, ancient sequences dated up to 60 ky are available for the mitochondrial D-loop, and could potentially be obtained for the nuclear genome, if the library improvements described above lead to sufficient genotype quality. However, several key evolutionary events are situated at intermediate ages between fossil and tip calibrations, and we lack a dedicated methodological framework to synergise and account for temporal information from both time periods, as pointed out in the thesis summary above.

With the bovid mitochondrial data used in Chapter 4, one attempt has been made to use the *BEAST framework to combine the implementation of the coalescent for intraspecific diversity (using aDNA as the calibration) and the implementation of a Yule prior for interspecific level (using the Yak/Bison fossil calibration; Figure C2).

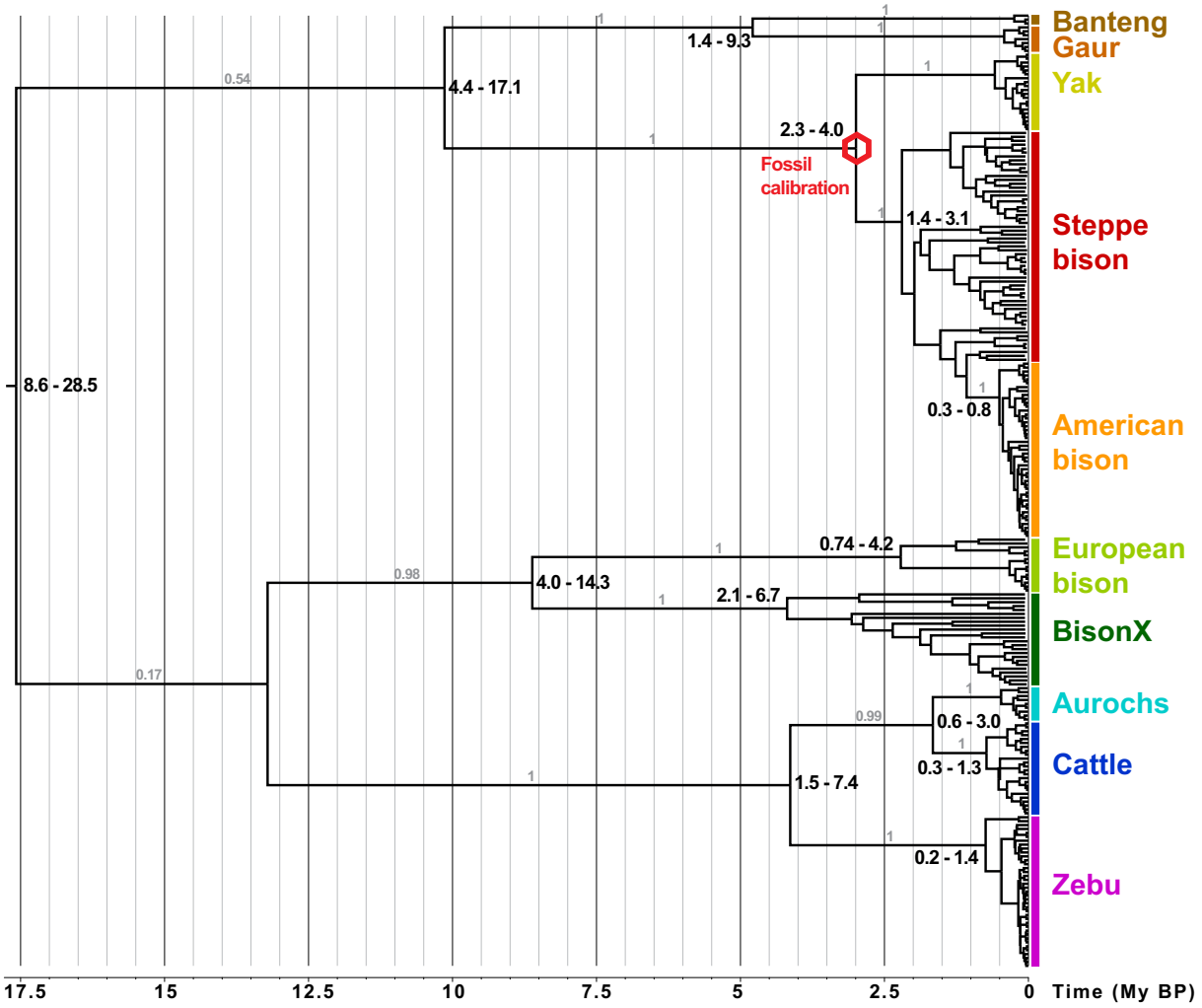


Figure C2. Maximum clade credibility tree from the Bayesian analysis of 611 bp D-loop sequences from 244 bovid modern and ancient samples. The analysis was conducted using the BEAST package (Drummond, Rambaut 2007). The substitution model TN93+G6 was selected through comparison of Bayesian Information Criterion in ModelGenerator v0.85 (Keane et al. 2006). The *BEAST option was used, defining each clade coloured on the tree as a separate population (at the exception of Steppe and American bison, combined into a single coalescent) and an uncorrelated relaxed clock was used for rate variations among lineages. The tree was calibrated by setting the divergence between Yak and Bison at around 2.5 million years (My), using a lognormal distribution with a mean of 2.5 My and 95% of the prior probability between 2 and 3 My (Wang et al. 2010).

Although the mitochondrial phylogeny of bovid is correctly recovered (see chapter 3), the inferred timescale is strongly overestimated. For example, the tMRCA of cattle is calculated at 0.76 (0.3 – 1.3) My, which is in complete contradiction with the timing of cattle domestication from archaeological evidence. Even species with a large number of ancient sequences, and hence more temporal information, present tMRCA estimates which are incompatible with inferences generated at the intra-specific level: for example 4.22 (2.1 – 6.7) My old tMRCA is calculated for BisonX, when coalescent calculations in Chapter 2 lead to an estimate of 151.7 (110.8 – 200.2) ky. It appears that the temporal information contained in the dates associated with ancient samples is not being adequately taken into account during the inference process, and instead the calculation largely relies on the fossil calibration point.

Another possibility to examine evolutionary events on both short and long timescales in the same phylogeny would be to perform individual coalescent calculations for each species, and then use the estimated tMRCA (including errors) as priors on the corresponding nodes of the *BEAST species tree. While this workaround is worth exploring, it seems that novel methodological developments will be required to accurately account for different types of calibration through time in the same phylogenetic analysis.

References

Drummond, AJ, A Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC evolutionary biology 7:214.

Drummond, AJ, MA Suchard, D Xie, A Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution.

Heled, J, AJ Drummond. 2010. Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution 27:570-580.

Ho, SYW, SO Kolokotronis, RG Allaby. 2007. Elevated substitution rates estimated from ancient DNA sequences. Biology Letters 3:702-705.

Ho, SYW, R Lanfear, MJ Phillips, I Barnes, JA Thomas, SO Kolokotronis, B Shapiro. 2011. Bayesian estimation of substitution rates from ancient DNA sequences with low information content. Systematic biology 60:366-375.

Keane, TM, CJ Creevey, MM Pentony, TJ Naughton, JO McLnerney. 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. BMC evolutionary biology 6:29.

Lanfear, R, B Calcott, SYW Ho, S Guindon. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution.

Posada, D. 2008. jModelTest: phylogenetic model averaging. Molecular Biology and Evolution 25:1253-1256.

Soares, P, L Ermini, N Thomson, M Mormina, T Rito, A Rohl, A Salas, S Oppenheimer, V Macaulay, MB Richards. 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. American Journal of Human Genetics 84:740-759.

Wang, Z, X Shen, B Liu, et al. 2010. Phylogeographical analyses of domestic and wild yaks based on mitochondrial DNA: new data and reappraisal. Journal of Biogeography 37:2332-2344.