Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes

M. Thomas P. Gilbert[a,b], Daniela I. Drautz[c], Arthur M. Lesk[c], Simon Y. W. Ho[d], Ji Qi[c], Aakrosh Ratan[c], Chih-Hao Hsu[c], Andrei Sher[e], Love Dalén[f], Anders Götherström[g], Lynn P. Tomsho[c], Snjezana Rendulic[c], Michael Packard[c], Paula F. Campos[a], Tatyana V. Kuznetsova[h], Fyodor Shidlovskiy[i], Alexei Tikhonov[j], Eske Willerslev[a], Paola Iacumin[k], Bernard Buigues[l], Per G. P. Ericson[m], Mietje Germonpré[n], Pavel Kosintsev[o], Vladimir Nikolaev[p], Malgosia Nowak-Kemp[q], James R. Knight[r], Gerard P. Irzyk[r], Clotilde S. Perbost[r], Karin M. Fredrikson[s], Timothy T. Harkins[s], Sharon Sheridan[s], Webb Miller[b,c], and Stephan C. Schuster[b,c]

[a]Centre for Ancient Genetics, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark; [c]Center for Comparative Genomics and Bioinformatics, Pennsylvania State University, 310 Wartik Building, University Park, PA 16802; [d]Centre for Macroevolution and Macroecology, School of Botany and Zoology, Australian National University, Canberra ACT 0200, Australia; [e]Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, 33 Leninsky Prospect, Moscow 119071, Russia; [f]Centro UCM-ISCIII de Evolución y Comportamiento Humanos, c/Sinesio Delgado 4, 28029 Madrid, Spain; [g]Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyv. 18D, S-752 36 Uppsala, Sweden; [h]Department of Paleontology, Faculty of Geology, Lomonosov Moscow State University, Leninskiye Gory, Moscow 119991, Russia; [i]The Ice Age Museum, All-Russia Exhibition Centre, Pavilion 71, Moscow 129223, Russia; [j]Zoological Institute, Russian Academy of Sciences, Universitetskaya nab. 1, St. Petersburg 199034, Russia; [k]Department of Earth Sciences, University of Parma, Parco Area delle Scienze 157/A, 43100 Parma, Italy; [l]2 Avenue de la Pelouse, F-94160 St. Mandé, France; [m]Department of Vertebrate Zoology, Swedish Museum of Natural History P.O. Box 50007, S-104 05 Stockholm, Sweden; [n]Department of Palaeontology, Royal Belgian Institute of Natural Sciences, Vautierstraat 29, 1000 Brussels, Belgium; [o]Institute of Plant and Animal Ecology, Urals Branch of the Russian Academy of Sciences, 202 8th of March Street, Ekaterinburg 620144, Russia; [p]Institute of Geography, Russian Academy of Science, Department of Glaciology, 29 Staromonetny per., Moscow 109017, Russia; [q]Oxford University Museum of Natural History, Parks Road, Oxford OX1 3PW, United Kingdom; [r]454 Life Sciences, 20 Commercial Street, Branford, CT 06405; and [s]Roche Diagnostics Corporation, 9115 Hague Road, Indianapolis, IN 46250-0414

We report five new complete mitochondrial DNA (mtDNA) genomes of Siberian woolly mammoth (*Mammuthus primigenius*), sequenced with up to 73-fold coverage from DNA extracted from hair shaft material. Three of the sequences present the first complete mtDNA genomes of mammoth clade II. Analysis of these and 13 recently published mtDNA genomes demonstrates the existence of two apparently sympatric mtDNA clades that exhibit high interclade divergence. The analytical power afforded by the analysis of the complete mtDNA genomes reveals a surprisingly ancient coalescence age of the two clades, ≈1–2 million years, depending on the calibration technique. Furthermore, statistical analysis of the temporal distribution of the ^{14}C ages of these and previously identified members of the two mammoth clades suggests that clade II went extinct before clade I. Modeling of protein structures failed to indicate any important functional difference between genomes belonging to the two clades, suggesting that the loss of clade II more likely is due to genetic drift than a selective sweep.

mtDNA genome | phylogeny | ancient DNA | next-generation sequencing

Although ancient DNA analyses offer the potential to tackle a tantalizing range of otherwise unapproachable questions, the actual achievements of the field have been limited by the postmortem degradation of DNA. Even in well preserved specimens from arctic environments, number of specimens and amount of data per specimen are limited. Previous studies to assess the genetic structure of extinct species, including mammoths (1), have had to rely on short sequence intervals that were often only a few hundred nucleotides in length. This has made it difficult to obtain precise estimates of substitution rates and divergence times, particularly for species exhibiting low levels of genetic variation. Additionally, it is possible that the accuracy of these estimates has been compromised by the presence of sequence damage in the form of miscoding lesions, which can introduce significant biases in estimates of evolutionary parameters (2). These problems can be addressed by large-scale sequencing with manifold coverage, which will increase the amount of informative data while filtering out the spurious polymorphisms resulting from sequence damage. This should serve to increase both the precision and accuracy of demographic estimates.

In this study, we have taken advantage of recent developments in high-throughput DNA sequencing to assemble one of the largest ancient mitochondrial DNA (mtDNA) datasets to date, consisting of a total of nearly 300,000 nucleotides of unique sequence data from 18 individual samples. By exploiting permafrost-preserved hair shaft material as a source of ancient DNA (3), we present five newly sequenced Siberian woolly mammoth mtDNA genomes (Fig. 1). In combination with the 13 previously published (3–7), these make it possible to scan for signs of natural selection along the mitochondrial genome and allow further investigation of the population structure discovered in past studies (1, 8), including the inference of a more precise evolutionary time scale. Analysis of the combined dataset indicates a deep temporal split between the two clades (I and II). This observation, coupled with statistical analysis of the temporal distribution of the ^{14}C ages of these and previously identified members of the two mammoth clades (1), suggests that, although they are apparently sympatric, clade II vanished from Siberia long before clade I.

Results and Discussion

Sequencing of Mitochondrial Genomes from Clade I and II Specimens. Using the recently published approach of adopting ancient hair shafts as a source of genetic material (3), we have generated five
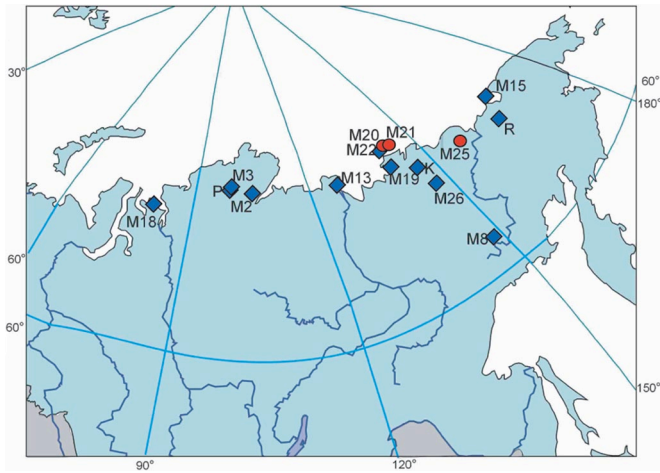
EVOLUTION

Fig. 1. Sites of recovery of the mammoth hair specimens whose complete mitochondrial genome sequences have been reported. Clade I mammoths are indicated as blue diamonds. Clade II mammoths are indicated by red circles. The exact locations of M1, M4, and M5 are not known but most probably originate from Northern Yakutia (c. 66–76°N, 106–160°E). ''K,'' ''R,'' and ''P'' indicate the Krause (4), Rogaev (5), and Poinar (6, 7) mammoth mtDNA genomes, respectively. This figure is modified from ref. 3.

novel mammoth mitochondrial genomes [Table 1 and supporting information (SI) Table S1]. The sequence data indicate that two of the sequences, M15 and M19, belong to clade I of the two mammoth clades recently identified (1). M19 is the well studied Yukagir specimen, whose good preservation state is manifested by long sequence read length and low DNA damage rate (Table 1). The geographic range of clade I mammoths is extremely large, spanning >6,000 km east to west and 20° in latitude, and encompassing both Siberian and North American specimens (1) (see Fig. S1). Both samples reported here were recovered from within this geographic range. In contrast, the geographic range reported previously for clade II mammoths is much more limited, spanning only 450 km (east to west) across the northern reaches of Siberia (1) (Fig. S1). Our data reveal that the three remaining sequences, M20, M21, and M25, are the first complete mitochondrial genomes of the second mammoth clade. This finding expands the observed range of clade II to ≈1,100 km (east to west), although it still appears to be limited to the region between the Lena and Kolyma rivers.

AMS^{14}C dating of the two new clade I individuals indicates that they are 13,995 ± 55 (M15) and 18,560 ± 50 (M19) ^{14}C years old. In contrast, the three clade II members were radiocarbon-dated as much older (Table 1). Only M25 had a finite ^{14}C age (59,300 ± 2,700 ^{14}C years before the present), whereas the ages of M20 and M21 were both beyond the limit of ^{14}C dating (M20 > 63,500 ^{14}C years; M21 > 58,000 ^{14}C years). This places the three clade II specimens as the oldest of the 18 mammoth mitochondrial genomes that have so far been reconstructed (see Table 1).

Comparative Analysis of 18 Mammoth Mitochondrial Genomes. The two mammoth clades are clearly observable within the 18 mammoth mtDNA genomes now available (Fig. 2, Fig. 3, Table 1, and Table S3). In addition to mammoths M15 and M19 (Yukagir), mammoth clade I also contains the two previously published mammoth mitochondrial genomes (4, 5) (here termed "Krause" and "Rogaev," respectively) and one assembled by us from previously published sequence data (6, 7) (here termed "Poinar"), as well as 10 published by our group in a recent study on ancient hair genomics (3). Among these 10 are the well known

Table 1. Description of the mammoth mitochondrial sequences

Sample	Tissue	^{14}C date	Year collected*	Sequencing technology	% mito-chondrial†	Contigs‡	Fold coverage	Average untrimmed read length§	% C→T damage	% trimmed read identity¶	% diff. vs. M1‖	% diff. vs. M25‖	% diff. vs. 591 bp M1**	% diff. vs. 591 bp M25**
M15	Hair	13,995 ± 55	n.d.	454/PCR	0.30	2	4.8	86.5	0.576	99.82	0.07	1.24	0.68	2,54
M19 "Yukagir"	Hair	18,560 ± 50	2003	454	1.86	1	72.7	199.8	0.202	99.88	0.15	1.22	0.68	1.86
M20††	Hair	>63,500	2000	454/PCR	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	1.14	0.24	2.88	0.34
M21	Hair	>58,000	2001	454	0.43	1	12.7	73.7	0.677	99.65	1.30	0.12	2.71	0.51
M25	Hair	59,300 ± 2,700	2003	454	1.43	1	25.4	113.4	0.712	99.64	1.26	0.00	2.54	0.00

*Year sample was discovered (where known).
†Percentage of mitochondrial sequences among sequence reads.
‡Number of contigs assembled out of mitochondrial sequences.
§Average read length before trimming based on Krause (4) sequence.
¶Average percent identity with respect to assembly after automatic computational quality processing (i.e., final read used in alignment).
‖Percent difference from M1 (clade I representative) and M25 (clade II representative) sequences, respectively.
**Percent difference over the 591-bp intervals used for mammoth–elephant comparison from M1 and M25, respectively.
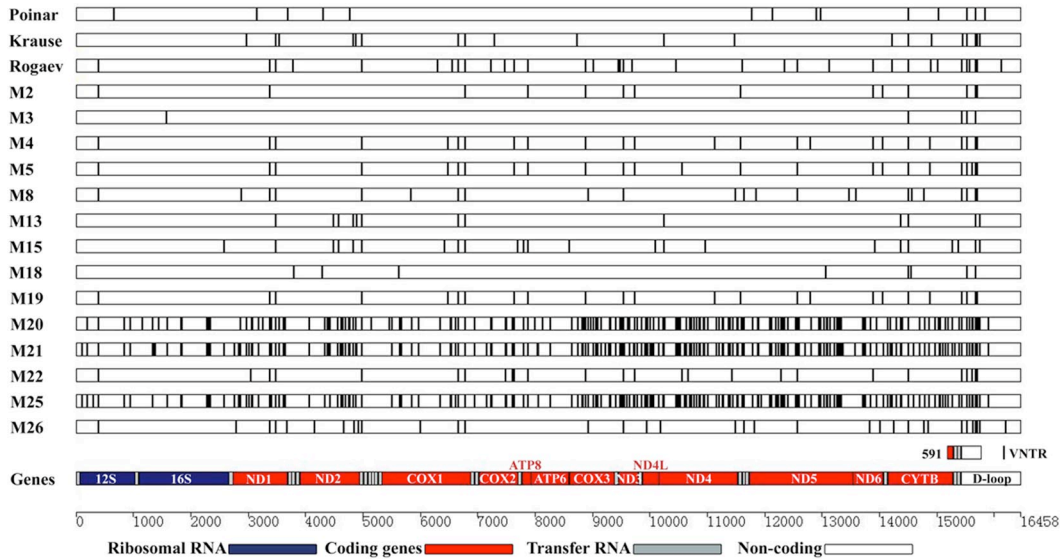††Summary statistics could not be calculated because of the low 454 coverage of sample M20.

Fig. 2. Sequence differences found among the 18 mammoth mitochondrial genomes with respect to mammoth M1 (GenBank entry EU153444.1). Each vertical bar depicts a nucleotide difference from sample M1, which serves as a reference (and hence has no row). The rectangle labeled 591 shows the location of a 591-bp interval used to assess the diversity among the larger mammoth and modern elephant datasets. We have not tried to assemble the interval denoted by VNTR; thus, this section is absent from the alignment.

Jarkov (M2), Fishhook (M3), Dima (M8), and Adams (M13) specimens.

Using a Bayesian phylogenetic method, we estimated the phylogeny and divergence times of several Proboscidean species (Fig. 3a). The chief difficulty in this divergence dating analysis was the selection of an appropriate calibration point. The fossil
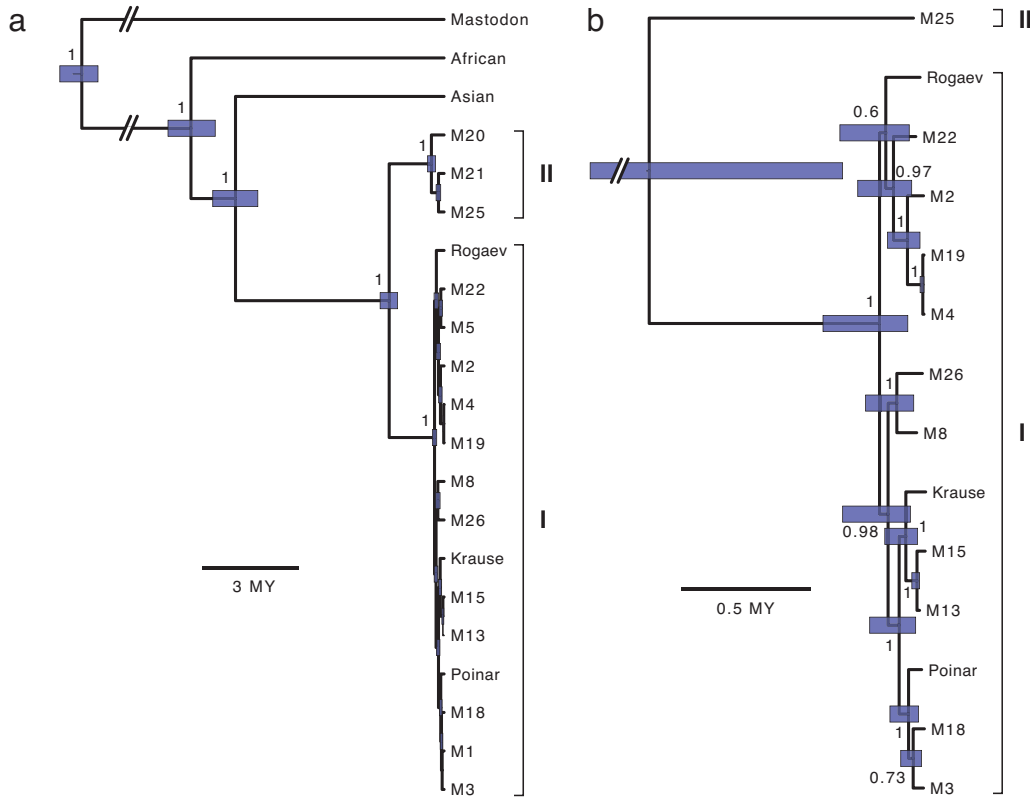
Fig. 3. Phylogenetic trees inferred using Bayesian analysis of complete mitochondrial genomes, drawn to time scales, with mammoth clades indicated. Nodes of interest are labeled with posterior probabilities. Blue bars represent 95% highest posterior densities of nodal age estimates. Slanted double lines indicate that portions of lines or bars have been omitted because of space constraints. (*a*) Estimated phylogeny of 18 mammoths, mastodon, and African and Asian elephants, where divergence dates are estimated using fossil calibration. (*b*) Estimated genealogy of 14 mammoth specimens with finite radiocarbon dates, where divergence dates are derived using an internally calibrated molecular clock.

Table 2. Nucleotide diversity within and between mammoth clades and elephant species over complete and partial (≈591 bp) mtDNA genomes

| | Mammoths | | | Elephants | | | |
	Clade I (Siberian*)	Clade II[†]	Interclade (Siberian*)	L. cyclotis[‡]	L. africana[‡]	Loxodonta (all)[§]	Elephas
Complete	0.0018	0.0011	0.012				
Partial	0.0091 (0.0043)	0.0061	0.0117 (0.0102)	0.0164	0.0292	0.0288	0.0177

*Calculated nucleotide diversity when dataset was restricted to Siberian mammoth dataset only.
[†]Clade II mammoths are geographically restricted to Arctic Siberia.
[‡]Nucleotide diversity calculated under the hypothesis that the two *Loxodonta* taxa represent two species (25).
[§]Nucleotide diversity calculated under the hypothesis that the two *Loxodonta* taxa represent one species (26).

record offers an age estimate for the divergence between mastodon and mammoth at ≈24–28 million years (MY), but this external calibration is possibly too deep for considering intraspecific divergences (9). An alternative is to analyze only the mammoth sequences, using their known ages as internal calibrations at the tips of the evolutionary tree (Fig. 3b). These calibrations, however, may be too shallow for investigating the deep interclade divergence. Therefore, we present estimates made by using both approaches, and suggest that the true dates lie between the two extremes.

By using the external, fossil-based calibration, the split between mammoth and Asian elephant was estimated at 6.45 MY, with a 95% highest posterior density (HPD) of 5.76–7.16 MY. This was preceded by the divergence between these two species and the African elephant, which occurred 7.83 MY ago (95% HPD: 7.08–8.54 MY). The estimated age of the African–Asian elephant separation is consistent with the 7.6 MY date inferred by Rohland *et al.* (10).

The timing of the coalescence between the two mammoth clades was estimated to be 1.70 MY ago (95% HPD: 1.44–1.98 MY) and 1.07 MY ago (95% HPD: 0.38–2.43 MY) by using external and internal calibrations, respectively. Together, these two date estimates suggest that the clade divergence occurred ≈1–2 MY ago.

Intraclade Nucleotide Diversity. The large number of differences observed between representative samples of the two clades (excluding the VNTR region, where data are absent) is in stark contrast to the low variation observed within each clade in this dataset (Table 2 and Fig. 2), as indicated by computation of nucleotide diversity (11). This, however, is likely to be an inaccurate representation of the true variation within and between the mammoth clades, because of the limited number of clade II samples in our complete mtDNA genome dataset and the absence of North American samples. To circumvent this limitation and perform an analysis of global mammoth nucleotide diversity that can be compared with modern elephant data, it is necessary to restrict the analysis to a 591-bp subsection of the 741-bp fragment sequenced by Barnes *et al.* (1), which covers 3.5% of the whole mitochondrial genome and includes part of the cytochrome *b* gene, two full tRNA genes, and 356 bp of the D-loop (see Fig. 2). Combined with the newly generated data, this yields an alignment of 47 clade I and 12 clade II mammoths, with representatives from Siberia and North America. Analysis of this expanded dataset increases the observed nucleotide diversity as expected, much of which is due to the presence of non-Siberian samples (Table 2). A comparison of these data with the nucleotide diversity of living elephants, using the homologous genetic regions from 97 African elephants (16 *Loxodonta cyclotis*, 81 *Loxodonta africana*) and 43 Asian elephants (*Elephas indicus*) that are present in GenBank, suggests that despite their large geographic range, mammoth mtDNA nucleotide diversity over this genetic region was considerably less than that observed in modern elephants (Table 2). This may be an effect of possible differences between mammoths and extant elephants in their population size, geographical range of sample collection, or even reproductive differences between the species. As further complete elephant and mammoth mtDNA genomes are sequenced it will be possible to discern whether this pattern holds true and to investigate these issues further.

Clade Distribution Through Space and Time. Phylogenetic analysis of the complete mtDNA genomes demonstrates the existence of two highly diverged mammoth clades that were sympatric in space and time (Fig. S1). It has previously been noted that clade I had a large distribution, throughout Beringia, during marine isotope stage 3 (MIS 3: 60–25 kya), whereas clade II seems to have been restricted to the region between the Lena and Kolyma rivers (1). However, although both clades coexisted in the latter region for thousands of years (1), the distribution of the ages of ^{14}C-dated mammoths suggests an extended presence of clade I in the paleontological record for tens of thousands of years after evidence for clade II ceases to exist (Table 1). This observation becomes more pronounced when combined with the ^{14}C dated samples published previously (1) (total dataset of 43 clade I and 11 clade II mammoths; see Fig. 4). We evaluated the likelihood that a constant ratio of the two clades existed side-by-side until a simultaneous extinction. The analysis suggested a very low probability of such a pattern arising by chance, given stable proportions of both clades ($P = 0.002$, based on a simulation with 100,000 permutations). When one of the clade II samples that could only be dated as >33,000 ^{14}C years (1) is removed from the analysis, the probability is much lower ($P = 0.0008$).

Selection or Drift? The observation of two clades coexisting for an extended period, followed by the extinction of one of them, raises a number of questions with regard to their evolutionary relationship. The presence of two very different mitochondrial genomes in Siberia is not reflected in morphological variation of *M. primigenius* as currently described, which provides no evidence of more than one species of mammoth coexisting in Siberia within the last 300,000 years (12). It therefore seems unlikely that the two clades are related to the existence and asynchronous extinction of two reproductively isolated groups of mammoths.

There are, however, several additional explanations for the observation of an extended clade I survival. One possibility is that mitochondrial genomes belonging to clade I had a selective advantage over those belonging to clade II. The extinction of clade II could thus be due to a selective sweep. The sequencing of heterochronous and complete mtDNA genomes allows for a unique possibility to address this hypothesis directly. To investigate potential functional genomic differences among the two clades, we assessed nonsynonymous substitutions in mitochondrial-encoded proteins, searching for amino acid replacements that could have influenced protein function (for details, see *SI Text* and Table S4). A total of 31 amino acid replacements were
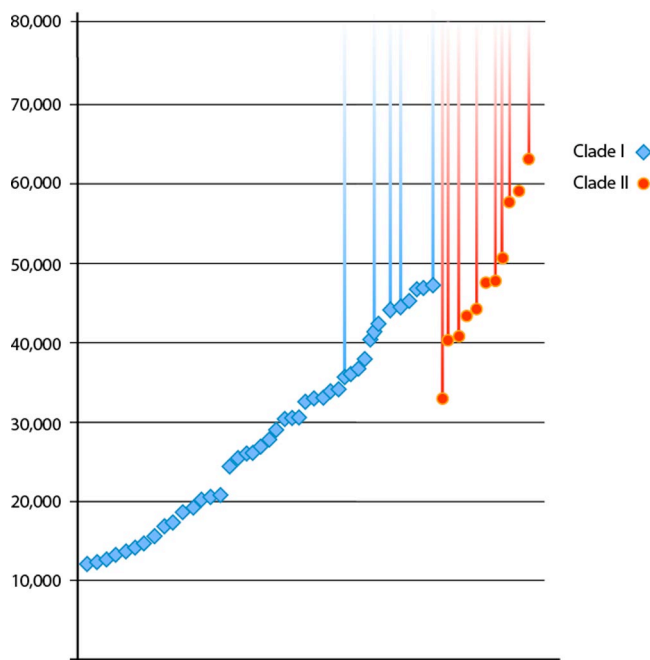
Fig. 4. Temporal distribution of 43 ^{14}C dated clade I and 11 clade II mammoths. For a number of individuals (5 clade I and 8 clade II, indicated by the extended timelines) finite dates could not be calculated; thus, the reported dates are only indicative of minimum bounds on the samples' ages. In this context, some aspects of the figure may be misleading—for example, the apparent absence of clade I mammoths between ≈50,000 and 60,000 ^{14}C years ago. The apparently ''older'' minimum on the infinite dates of most of the hair (this study) as compared with bone samples (1) may reflect a superiority in hair over bone material with regard to long-term carbon preservation in the samples.

discovered. Moreover, clade II mammoths and Asian/African elephants have five and four additional residues, respectively, on the C-terminal end of the *ND4* gene compared with the clade I mitochondrial genome. All of the observed substitutions appear to be between closely related amino acids. For those proteins having a close homolog with an experimentally determined structure (namely, *COX1*, *COX2*, *COX3*, and *Cytb*), we also modeled the structure of the mammoth proteins. All substitutions appear in regions on the surface or in loop regions that neither seem essential for proper folding nor would be expected to alter protein function in any obvious way (see *SI Text*). Therefore, the evidence from the modeled structures suggest that it is unlikely that the nonsynonymous differences found in the mitochondrial genomes of the two mammoth clades have resulted in any physiological disparities, and thus a selective advantage for clade I based on mtDNA sequence differences alone is not expected.

A more likely alternative is that the loss of clade II is a consequence of its restricted geographical distribution, because taxa with small ranges are generally more prone to extinction compared with widespread taxa (13). It is therefore conceivable that clade II was lost because of a demographic bottleneck resulting in genetic drift or a local population extinction. Taking into account the previous observation of an overall large and stable population size in mammoth during MIS 3 (1), we hypothesize that this population decline or extinction was limited to the Lena–Kolyma region of Siberia. Additional sampling in this region would make it possible to test this hypothesis further and to better resolve the timing of the loss of clade II.

Conclusions

Our report of the first complete clade II mtDNA genomes in Siberian mammoths offers tantalizing insights into the history of this iconic species. Although no functional differences were observed between the different mtDNA genomes, the deep phylogenetic split between the two clades and their apparent coexistence, in combination with statistical evidence that indicates the demise of clade II up to 30,000 thousand years before the demise of clade I on mainland Siberia, raises questions about the evolutionary relationship between the two clades. On the one hand, the data could simply represent natural variation within a single species, driven by an early maternal lineage split that was retained in later history, with the clade II sequences disappearing because of genetic drift. A number of extant mammalian taxa show similar patterns of genetic variation within single continuous populations, including moose (14), reindeer (15), and Asian elephants (16), and several ancient DNA studies have shown similar patterns of clade extinction in taxa such as cave bears (17), brown bears (18), wolves (19), and arctic fox (20). On the other hand, in light of the incompleteness of most fossil mammoth remains and our inability to make observations on living mammoths to provide behavioral or other cues that aid in the resolution of the taxonomic relationship between living mammals, the genetic data may provide evidence of something more, specifically the existence of sympatric mammoth species that underwent asynchronous extinction events.

As further mammoth complete mtDNA genomes become available, including samples that represent the western and eastern limits of the range of the mammoth, and as nuclear DNA analyses are applied to representatives of the two clades, it will become possible to test three explicit hypotheses. (*i*) The observed phylogenetic divergence represents a relict ancestral polymorphism that has formed without the existence of any barriers to gene flow (21). This would be consistent with the fact that northeast Siberia is the "core area" for woolly mammoth, where the species was continuously present (and probably with substantial population size) through the Middle and Late Pleistocene (22). (*ii*) The genetic structure represents a mixing of populations that had evolved in isolation, for example, on either side of the Bering Strait before an admixture during the last glaciation (1). (*iii*) The material covers a wider range of time and includes earlier and later, genetically different mammoths. It might naturally be of interest, with regard to this particular point, for future studies to renew morphometric analyses on mtDNA profiled mammoth specimens, to investigate whether subtle phenotypic differences might be identified. By testing these three hypotheses, we will come closer to understanding the biology and extinction of the mammoth.

Materials and Methods

Complete mtDNA Sequence Generation. The mtDNA sequences were extracted from hair shaft samples, sequenced, and assembled following the procedure described in ref. 3, with the exception of samples M15 and M20, which were principally assembled from FLX generated sequences but completed by using conventional PCR approaches. The DNA sequences have been deposited in GenBank. As with our previous study (3), we have not assembled the VNTR region because of its extreme sequence variability. For details on the sample sources, geographic origins, and materials used, see Tables S1 and S3. Because of the extreme divergence of the clade II mammoths from those in the clade I, resequencing was performed to confirm the identity of the nucleotide positions that the 454 sequencing identified as divergent between M25 and the Krause (4) mammoth genome. Details of the primers and regions amplified can be found in Table S2. PCR was performed for 40 cycles, using High Fidelity Platinum Taq (Invitrogen) at Pennsylvania State University, in 25-μl reactions according to the manufacturer's guidelines. Amplified DNA was directly sequenced in both directions by using Applied Biosystems BigDye sequencing chemistry on an ABI 3100.

[14]C Dating. Hair shaft samples were submitted from mammoths M15, M20, M21, and M25 to the commercial ^{14}C dating facility at the University of Oxford. For ^{14}C sample identification numbers, see Table S1.

EVOLUTION

Phylogenetic Analysis. Phylogenetic analysis of *Mammuthus*, *Elephas*, *Loxodonta*, and *Mammut* was performed with the program BEAST 1.4.6 (23). Eighteen mammoth mitochondrial genomes from this and previously published (3–7) studies were used in combination with Asian elephant (GenBank entry NC_005129) (5), African elephant (NC_000934) (24), and mastodon (EF632344) (10). The TrN+I model of nucleotide substitution was used, as selected by comparison of Akaike information criterion values. Two separate phylogenetic analyses were performed: (*i*) analysis of the complete dataset, calibrated by using a lognormal prior on the age of the mammoth–mastodon divergence (minimum 24 MY, mean 26 MY, with 95% of the distribution lying between 24 and 28 MY), with a constant-size coalescent prior on the mammoth clade; and (*ii*) analysis of the 14 mammoth genomes with finite radiocarbon dates, using their known ages as calibrations on the tips of the tree, with a constant-size coalescent prior on the entire tree.

In both cases, posterior distributions were obtained by Markov chain Monte Carlo (MCMC) sampling. Samples were drawn every 1,000 MCMC steps from a total of 2,000,000 steps, following a discarded burn-in of 200,000 steps. Acceptable mixing and convergence to the stationary distribution were checked by inspection and plotting of posterior samples.

Statistical Analysis of Mammoth Clade Extinction. The statistical test of the temporal distribution of the clade I and II mammoth remains used the complete ^{14}C dated dataset of this study and that reported by Barnes *et al.* (1). Among the $N = 59$ dated specimens, the 11 clade II mammoths have frequency $F = 11/59 = 0.1864$. Let M be the number of samples (either clade) that are at least as old as the youngest clade II sample (33,000 years). In our case, $M = 27$ if one takes the Barnes data at face value. We wish to test the possibility that the absence of recent clade II individuals is due to sampling error; specifically, we want to reject the hypothesis that clade II existed side-by-side with clade

I at constant frequency F up to a simultaneous extinction. Informally, if we generate N random positive integers, assigning each to clade I with probability $1 - F$ or to clade II with probability F, how frequently will all clade II assignments be among the first M? To generate empirical P-values, we analyzed 100,000 random sequences of 59 numbers. In 215 cases, all of the assignments to clade II occurred within the first $M = 27$ of the $N = 59$ numbers ($P = 0.00215$). Removing the sample of Barnes *et al.* with putative (and likely incorrect) age of 33,000 years gives $N = 58$, $F = 10/58 = 0.1724$, $M = 19$, and an empirical P-value of $83/100,000 = 0.00083$ that the extreme age skew of the clade II samples occurred by chance. The results of the statistical test are necessarily conservative because the ^{14}C dates of a number of the clade II mammoths were beyond the ^{14}C dating limit (Fig. 4); for these, the test was performed only on the sample-specific ^{14}C limit (which could potentially be many thousands of years closer to the present than the true sample age).

1. Barnes I, *et al.* (2007) Genetic structure and extinction of the woolly mammoth, Mammuthus primigenius. *Curr Biol* 17:1072–1075.
2. Ho SYW, Heupink TH, Rambaut A, Shapiro B (2007) Bayesian estimation of sequence damage in ancient DNA. *Mol Biol Evol* 24:1416–1422.
3. Gilbert MTP, *et al.* (2007) Whole-Genome shotgun sequencing of mitochondria from ancient hair shafts. *Science* 317:1927–1930.
4. Krause J, *et al.* (2006) Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. *Nature* 439:724–727.
5. Rogaev EI, *et al.* (2006) Complete mitochondrial genome and phylogeny of Pleistocene mammoth Mammuthus primigenius. *PLoS Biol* 4:e73.
6. Poinar HN, *et al.* (2006) Metagenomics to palaeogenomics: Large-scale sequencing of mammoth DNA. *Science* 311:392–394.
7. Gilbert MTP, *et al.* (2007) Recharacterization of ancient DNA miscoding lesions: Insights in the era of sequencing-by-synthesis. *Nucleic Acids Res* 35:1–10.
8. Höss M, Pääbo S, Vereshchagin NK (1994) Mammoth DNA sequences. *Nature* 370:333.
9. Ho SYW, Larson G (2006) When times are a-changin'. *Trends Genet* 22:79–83.
10. Rohland N, *et al.* (2007) Proboscidean mitogenomics: Chronology and mode of elephant evolution using mastodon as outgroup. *PLoS Biol* 5:e207.
11. Nei M (1987) *Molecular Evolutionary Genetics* (Columbia Univ Press, New York).
12. Lister AM, Sher AV, van Essen H, Wei G (2005) The pattern and process of mammoth evolution in Eurasia. *Quat Int* 126–128:49–64.
13. Payne JL, Finnegan S (2007) The effect of geographic range on extinction risk during background and mass extinction. *Proc Natl Acad Sci USA* 104:10506–10511.
14. Hundertmark KJ, *et al.* (2002) Mitochondrial phylogeography of moose (Alces alces): Late Pleistocene divergence and population expansion. *Mol Phylogenet Evol* 3:375–387.
15. Flagstad Ø, Røed KH (2003) Refugial origins of reindeer (Rangifer tarandus L.) inferred from mitochondrial DNA sequences. *Evolution (Lawrence, Kans)* 57:658–670.
16. Fleischer RC, Perry EA, Muralidharan K, Stevens EE, Wemmer CM (2007) Phyologeography of the Asian Elephant (Elephas maximus) based on mitochondrial DNA. *Evolution (Lawrence, Kans)* 55:1882–1892.
17. Hofreiter M, *et al.* (2007) Sudden replacement of cave bear mitochondrial DNA in the late Pleistocene. *Curr Biol* 17:R122–R123.
18. Barnes I, Matheus P, Shapiro B, Jensen D, Cooper A (2002) Dynamics of Pleistocene population extinctions in Beringian brown bears. *Science* 295:2267–2270.
19. Leonard JA, *et al.* (2007) Megafaunal extinctions and the disappearance of a specialized wolf ecomorph. *Curr Biol* 17:1146–1150.
20. Dalén L, *et al.* (2007) Ancient DNA reveals lack of postglacial habitat tracking in the arctic fox. *Proc Natl Acad Sci USA* 104:6726–6729.
21. Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution (Lawrence, Kans)* 56:2383–2394.
22. Sher AV, Kuzmina S, Kiselyov S, Lister A (2003) Tundra-Steppe environment in Arctic Siberia and the evolution of the wooly mammoth. *Program and Abstracts. Occasional Papers in Earth Sciences no. 5, Paleontology Program*, ed Storer JE (Government of the Yukon, Canada), pp 136–142.
23. Drummond AJ, Rambaut A (2007) A BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.
24. Hauf J, Waddell PJ, Chalwatzis N, Joger U, Zimmermann FK (2000) The complete mitochondrial genome sequence of the African elephant (Loxodonta africana), phylogenetic relationships of Proboscidea to other mammals, and D-loop heteroplasmy. *Zoology* 102:184–195.
25. Roca AL, Georgiadis N, Pecon-Slattery J, O'Brien SJ (2001) Genetic evidence for two species of elephants in Africa. *Science* 293:1473–1477.
26. Debruyne R (2005) A case study of apparent conflict between molecular phylogenies: The interrelationships of African elephants. *Cladistics* 21:31–50.