Current Biology *17*, 1072–1075, June 19, 2007 DOI 10.1016/j.cub.2007.05.035

Report

Genetic Structure and Extinction of the Woolly Mammoth, *Mammuthus primigenius*

Ian Barnes,[1] Beth Shapiro,[2] Adrian Lister,[3,4,*] Tatiana Kuznetsova,[5] Andrei Sher,[6] Dale Guthrie,[7] and Mark G. Thomas[3]

[1] School of Biological Sciences
Royal Holloway
University of London
Egham, Surrey TW20 0EX
United Kingdom

[2] Department of Zoology
University of Oxford
Oxford OX1 3PS
United Kingdom

[3] Department of Biology
University College London
London WC1E 6BT
United Kingdom

[4] Department of Palaeontology
The Natural History Museum
Cromwell Road
London SW7 5BD
United Kingdom

[5] Department of Paleontology
Moscow State University
Moscow 119992
Russia

[6] Severtsov Institute of Ecology and Evolution
Russian Academy of Sciences
Moscow 119071
Russia

[7] Institute of Arctic Biology
University of Alaska
Fairbanks, Alaska 99709

Summary

The interval since circa 50 Ka has been a period of significant species extinctions among the large mammal fauna. However, the relative roles of an increasing human presence and a synchronous series of complex environmental changes in these extinctions have yet to be fully resolved [1]. Recent analyses of fossil material from Beringia have clarified our understanding of the spatiotemporal pattern of Late Pleistocene extinctions, identifying periods of population turnover well before the last glacial maximum (LGM: circa 21 Ka) or subsequent human expansion [2–4]. To examine the role of pre-LGM population changes in shaping the genetic structure of an extinct species, we analyzed the mitochondrial DNA of woolly mammoths in western Beringia and across its range. We identify genetic signatures of a range expansion of mammoths, from eastern to western Beringia, after the last interglacial (circa 125 Ka), and then an extended period during which demographic inference indicates no population-size increase. The most marked change in diversity at this time is the loss of one of two major mitochondrial lineages.

Results and Discussion

To examine changes in genetic diversity through time, we conducted DNA extractions of 96 bone, tooth, and ivory specimens with the appropriate protocols for working with ancient material (see Experimental Procedures). All ancient-DNA extractions and polymerase chain reaction (PCR) amplifications were conducted in a specialist ancient-DNA laboratory that was physically separated from post-PCR work. Overlapping segments of the mitochondrial genome were PCR amplified and sequenced to provide data, including the 3′ end of the cytochrome *b* gene, the threonine and proline tRNAs, and the first hypervariable portion of the control region. Analysis of 741 base pairs (bp) of mitochondrial DNA was possible for 41 mammoths, including individuals from Europe, Asia, and North America (see the Supplemental Data available online for specimen details) with radiocarbon dates ranging from before 50 Ka to 12 Ka (Figure 1). Of these, 33 were collected from western Beringia (Figure 1, sites D, G, and H), three from eastern Beringia (sites A and B), one from Kamchatka (site C), three from Central North Siberia (sites I, J, and K), and one from Europe (site L). We supplemented these data with two published western-Beringian sequences (sites E and F [5, 6]) and one from Central North Siberia (site I [7, 8]).

Bayesian phylogenetic analyses identified two well-supported, and previously unidentified, major mitochondrial lineages of mammoths in western Beringia and a third potential lineage in Europe (Figure 2). Because only one European specimen was included in this analysis, however, further work is in progress to confirm the existence of this third lineage.

We estimated the rate of evolution for this region of the mammoth mitochondrial genome with the radiocarbon-dated specimens as calibration points (Figure 2; see Experimental Procedures for details). Analyses were performed both with all specimens associated with finite radiocarbon dates and with a data set restricted to only western-Beringian specimens (localities D–H) with finite radiocarbon dates, so as to account for potential problems associated with geographic population structure. A mean rate estimate of 24% per My (95% highest posterior density [HPD], 6.3%–45%) was obtained. With this rate, dates and confidence intervals were estimated for the deeper nodes in the tree.

Our analyses provide a much more detailed picture of the later stages of the history of mammoth populations than is evident from the paleontological record alone [9, 10]. Examination of the most-well-represented lineage, clade 1, shows that sequences from our eastern-most sampling localities, in Alaska and Kamchatka, lie both at the base of clade 1 and within it. The presence

*Correspondence: a.lister@nhm.ac.uk

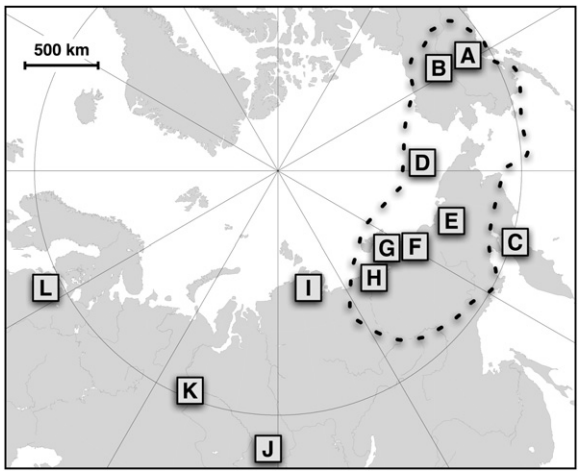

Figure 1. A Polar Projection Showing the Locations from which Samples Were Obtained

The dashed line indicates the approximate extent of Beringia.

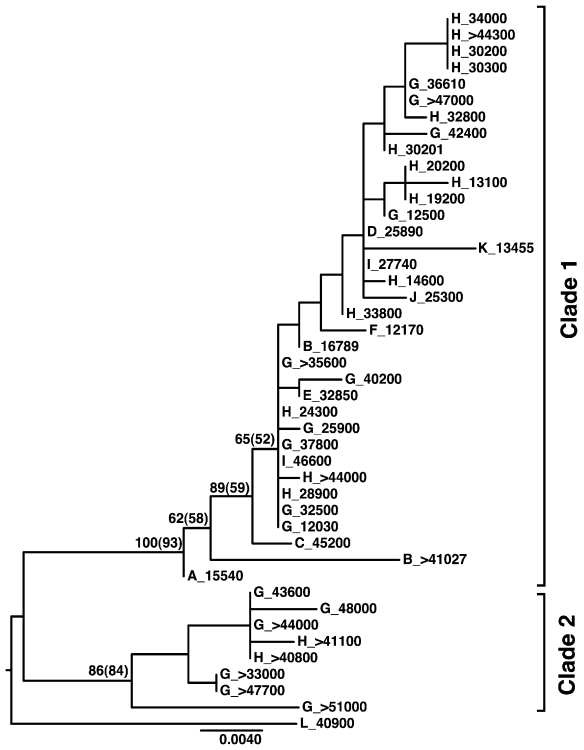

Figure 2. A Maximum-Likelihood Phylogeny for the 44 Mammoth Sequences Employed in This Study

Numbers to the left of nodes indicate the degree of support for those nodes on the basis of posterior probability and bootstrapping (in brackets). On branch tips, letter codes and colors indicate sample locations (as in Figure 1), and numbers indicate the point-estimate radiocarbon date. Full details of sequence analysis and sample provenance are given in the Experimental Procedures and Supplemental Data.

of clade 1 sequences from eastern Beringia that root at both ancestral and more derived nodes and the reduced diversity observed across Siberia suggest that the origins of this lineage are in the east. A lower bound for the timing of the western expansion of clade 1, taken as the time to the most recent common ancestor of the Siberian component of this clade (Figure S1, node A), is estimated at 63.4 Ka (HPD 47–94 Ka). This timing places the migration of mammoths after the last interglacial—marine isotope stage (MIS) 5e. Two factors would have restricted migration between eastern and western Beringia until after this time: The interglacial rise in sea levels would have submerged the Bering land bridge until at least around 100 Ka [11], and warmer climatic conditions would have led to northern expansion of boreal forest [12–14], constraining steppe-adapted mammoths to the far north.

After this period of post-MIS 5 western expansion, clade 1 can be observed across the sampling area, from Alaska to Central Siberia (Figures 1 and 2). Recovery of clade 2 sequences is restricted to the more heavily sampled western-Beringian sites (localities G and H). Interestingly, although both clades are present prior to the upper limit of radiocarbon dating (circa 50 Ka), only clade 1 persists until the end of the Pleistocene. Clade 2 is last observed at 43.6 Ka, indicating an at least local extinction of this lineage around that time.

Although the extinction of clade 2 might be attributed to selection against individuals with this haplotype, this seems unlikely. In this scenario, selection would require either that the clade 1 mitochondrial genome possessed some fitness advantage over clade 2, a situation rarely identified in wild populations [15], or that adaptive differences between the two populations had been maintained by reproductive isolation despite their sympatric distribution over millennia. A much more likely explanation is that clade 1 became fixed within the western-Beringian population by genetic drift under conditions of zero or negative population growth. An interpretation of constant population size is further supported by the regional stability of climate and vegetation during MIS 3 [16] and by Bayesian estimation of population-size changes for western-Beringian mammoths (Table 1 and Supplemental Data). For these analyses, we used the 33 finite dated sequences in phylogenetic analyses assuming two different fixed coalescent models of population history, (1) constant population size and (2)

Table 1. Exponential Growth Rates Estimated in Strict- and Relaxed-Clock BEAST Analyses of 33 Mammoth Sequences Associated with Finite Radiocarbon Dates

	Strict Clock (HKY+G)	Strict Clock (HKY+I)	Relaxed Clock (HKY+G)	Relaxed Clock (HKY+I)
Mean	6.62E-06	1.48E-06	3.90E-05	5.71E-05
95% HPD lower	−6.58E-06	−1.37E-05	−7.20E-06	−1.15E-05
95% HPD upper	2.23E-05	1.77E-05	1.10E-04	1.65E-04
Effective sample size	5267.414	5697.028	312.595	192.951

A growth rate of 0 was included in the posterior distribution of each analysis.

exponential growth or decline, and a more complex model in which the population history is estimated from the data across the sampling period (the Bayesian skyline plot [17]). In all three models, tree topology and substitution parameters were estimated simultaneously with the demographic-model parameters.

These analyses indicate a relatively constant population size throughout the time to the common ancestor of the sampled mammoths. Comparison of the mean (ln) posterior and mean (ln) coalescent likelihoods reveals that the data do not support a model of recent growth or decline. Although the exponential-growth model does appear to fit the data slightly better than the other models, the growth rate estimated from this analysis is extremely low, and in all four of the variants of this model examined (Table 1), a value of 0 (no growth) is included in the 95% confidence intervals.

Although the results of all three demographic analyses indicate a lack of population expansion or decline during MIS 3 (60–25 Ka), we note that the power of the Bayesian skyline plot to detect rapid decline and recovery from our data set of western-Beringian mammoths is weak: The signal for the loss of clade 2 during this period is not recovered, most likely because of the small number and relatively old dates of the two individuals representing this clade. However, because fine-scale sampling across this potential bottleneck is practically impossible because of the limits of finite radiocarbon dating, small and rapid changes in population size are unlikely to be detected in the genetic data by coalescent-based inference [18]. This might explain the apparent discrepancy between the estimate of population-size change based on the frequency of radiocarbon-dated material [16] and that based on the Bayesian analyses. Importantly, however, we can reject an MIS 3 population crash or local extinction of the type observed in bison, brown bears, and horses [2–4] in favor of an overall trend for population stability during MIS 3.

Recent research on the extinction of the mammoth fauna of northern latitudes in the Late Pleistocene and Holocene (circa12 Ka to the present) has defined dates and areas of final occurrence for several Holarctic species [3, 10, 19, 20]. These studies have highlighted the importance of environmental (particularly vegetational) changes in understanding the extinction process while also casting doubt on a single, simple explanation for Late Pleistocene extinctions (see also [1]). Ancient-DNA analyses have focused on past faunal migration, replacement, and changes in population size [2, 4], drawing attention to demographic events occurring prior to the onset of full-glacial conditions, during MIS 3. The local extinction of brown bears and stilt-legged horses in Alaska at circa 35 Ka [2, 3] and a decline in bison population size starting around 30–35 Ka [4, 17] hint at the importance of early, preglacial events in determining the late-glacial and postglacial status of these surviving taxa. However, the western-Beringian mammoths, in contrast to the other species so far studied, do not follow the pattern of complex population-size change and turnover during MIS 3. This emphasizes the importance of determining multiple individual-species histories in understanding the Late Pleistocene extinctions. We identify two phenomena that would have resulted in a more gradual loss of genetic diversity for this population: a founder effect following a post-MIS 5 expansion, and negligible or negative population growth during MIS 3. The extent to which this sample is representative of mammoths across their range and the relationship between demographic history and extinction remain to be fully determined.

Experimental Procedures

DNA was extracted with the method described in [2], modified by the collection of bone powder by the use of either a Spex freezer mill or a 20 mm drill bit at very low speed (approximately 60 rpm). Extractions were conducted in batches of 12 samples with at least one negative extraction control in each batch, with 100–200 mg of bone powder used in each extraction and a final extract volume of 30–50 ul.

PCR amplifications were performed in 25 μL reactions with 1 μL of extract, 1.25 U Platinum Taq Hi-Fidelity and 1X buffer (Invitrogen, UK), 2 mg/ml bovine serum albumin (BSA), 2 mM $MgSO_4$, 250 μM of each deoxynucleotide triphosphate (dNTP), and 1 μM of each primer. PCR thermal-cycling reactions were 94°C for 3 min, 40 cycles of 94°C denaturation for 30 s, annealing at 50°C –55°C (depending on the primer) for 45 s, and extension at 68°C for 45 s. Negative controls were used in each amplification, and on the rare occasions (n = 2) that these were positive (i.e., contaminated), the results of that set of amplifications were rejected. The 741 bp sequence contigs used in this study were generated with one of three amplification strategies, depending on the success of preliminary amplifications (Table S2). The resulting contigs cover the terminal 244 bp of cytochrome *b*, the complete tRNA-Thr and tRNA-Pro, and the first hypervariable portion of the control region. The overlapping nature of these PCR strategies results in a substantial degree of sequence replication for each sample (for strategy 1, 33% of base positions are replicated in a separate sequence, for strategy 2, 35%, and for strategy 3, 25%).

Sequencing reactions were conducted with the ABI PRISM Big Dye v2.0 chemistry and resolved with an ABI 3100 sequencer. A subset of amplifications made on four templates with the primer combination 15393F/15780R were cloned with Invitrogen's TOPO-TA cloning vector, and individual clones were sequenced to give some indication of template quality and the possible presence of nuclear copies. The absence of any clearly identifiable alternative sequence among these clones, coupled with the absence of any mismatch when the heavily overlapping primers in our amplification strategies were used, suggests that nuclear-encoded copies of mitochondrial sequences (numts) were not recovered. Sequences were found to be totally consistent between fragments generated by different primer pairs and replicable between amplifications when the same primer pairs were used (details available upon request).

Phylogenetic relationships among the 44 mammoths for which DNA-sequence information was available were estimated with the phylogenetic-analysis software MrBayes v3.1.2 [21] and PAUP v4b10 [22]. Substitution models were compared with likelihood-ratio tests (when nested) and the Akaike Information Criterion. The Hasegawa-Kishino-Yano + gamma + invariant sites (HKY+G+I) model, which incorporates different rates for transitions and transversions, rate variation across sites, and a proportion of invariable sites, was used to generate Bayesian posterior probabilities. Markov-chain Monte Carlo (MCMC) sampling was performed as implemented in MrBayes with the default settings (two runs of four chains each) for 10,000,000 iterations, with the first 10% discarded as burn in. Mixing and convergence to stationary distributions were investigated with Tracer v1.4 [23]. A maximum likelihood (ML) phylogeny was estimated with PAUP with the evolutionary model described above. An initial tree was generated by neighbor joining (NJ), followed by estimation of substitution-model parameters. A full heuristic search was then performed with subtree pruning and regrafting (SPR)-branch swapping with the parameters fixed to the estimated values. Parameters were then re-estimated, and the heuristic search was performed again. To determine support for the nodes, we performed 1000 full-heuristic bootstrap replicates with resampling, with starting trees generated by NJ and nearest-neighbor interchange (NNI)-branch swapping and with model parameters fixed to the estimated values.

To estimate a rate of evolution and dates of divergence between the clades, we performed both strict-clock and uncorrelated log-normal relaxed-clock analyses [24] as implemented in the software BEAST v1.4 [25] on a subset of the full data set consisting of 33 samples with finite radiocarbon dates. For both the strict- and relaxed-clock analyses, we used the HKY+I and the HKY+G model of nucleotide substitution, with sequence ages fixed to the mean uncalibrated radiocarbon date. For each analysis, the demographic function was estimated under different models (see Results and Discussion), including the Bayesian skyline plot [17] with five steps. Posterior distributions of parameters were approximated by MCMC sampling, with samples drawn every 3000 iterations over a total of 30,000,000 iterations after a discarded burn in of 10%. To check for consistency among the posterior distributions, we ran each analysis, and the results were combined. Mixing and convergence to the stationary distribution were evaluated, and the Bayesian skyline plot was calculated with Tracer v1.4. Posterior estimates for rate and divergence-date estimates were broadly similar between the analyses, although as expected the 95% HPD was wider when a relaxed clock was used (Table S3, Supplemental Data). Values reported in the main text are those from the best-fitting model (relaxed clock, HKY+I). To estimate the demographic history of the Siberian mammoth population, we performed molecular-clock analyses with the HKY+I model as above, but we limited the data set to 25 samples from western Beringia (sites D–H) with finite radiocarbon dates.

Supplemental Data

Experimental Procedures, two figures, and three tables are available at http://www.current-biology.com/cgi/content/full/17/12/1072/DC1/.

Acknowledgments

We thank Tony Stuart and Lembi Loũgas for sample provision, and Andrew Rambaut for assistance with data analysis. This work was supported by the United Kingdom Natural Environment Research Council (I.B.), the Royal Society (B.S.), and the Russian Foundation for Basic Research (A.S.). T.K. and A.S. thank the Russian-German Laptev Sea System program for contributing to the collection and dating of material.

Received: March 20, 2007
Revised: May 10, 2007
Accepted: May 15, 2007
Published online: June 7, 2007

References

1. Barnosky, A., Koch, P., Feranec, R., Wing, S., and Shabel, A. (2004). Assessing the causes of late Pleistocene extinctions on the continents. Science *306*, 70–75.
2. Barnes, I., Matheus, P., Shapiro, B., Jensen, D., and Cooper, A. (2002). Dynamics of Pleistocene population extinctions in Beringian brown bears. Science *295*, 2267–2270.
3. Guthrie, R.D. (2003). Rapid body size decline in Alaskan Pleistocene horses before extinction. Nature *426*, 169–171.
4. Shapiro, B., Drummond, A., Rambaut, A., Wilson, M., Matheus, P., Sher, A., Pybus, O., Gilbert, M., Barnes, I., Binladen, J., et al. (2004). Rise and fall of the Beringian steppe bison. Science *306*, 1561–1565.
5. Krause, J., Dear, P.H., Pollack, J.L., Slatkin, M., Spriggs, H., Barnes, I., Lister, A.M., Ebersberger, I., Paabo, S., and Hofreiter, M. (2006). Multiplex amplification of the mammoth mitochondrial genome and the evolution of *Elephantidae*. Nature *439*, 742–747.
6. Rogaev, E.I., Moliaka, Y.K., Malyarchuk, B.A., Kondrashov, F.A., Derenko, M.V., Chumakov, I., and Grigorenko, A.P. (2006). Complete mitochondrial genome and phylogeny of Pleistocene mammoth *Mammuthus primigenius*. PLoS Biol. *4*, e73.
7. Gilbert, M.T.P., Binladen, J., Miller, W., Wiuf, C., Willerslev, E., Poinar, H., Carlson, J.E., Leebens-Mack, J.H., and Schuster, S.C. (2007). Recharacterization of ancient DNA miscoding lesions: Insights in the era of sequencing-by-synthesis. Nucleic Acids Res. *35*, 1–10.
8. Poinar, H., Schwarz, C., Qi, J., Shapiro, B., MacPhee, R.D.E., Buigues, B., Tikhonov, A., Huson, D.H., Tomsho, L.P., Auch, A., et al. (2006). Metagenomics to paleogenomics: Large-scale sequencing of mammoth DNA. Science *311*, 392–394.
9. Lister, A., and Sher, A.V. (2001). The origin and evolution of woolly mammoth. Science *294*, 1094–1097.
10. Stuart, A.J., Kosintsev, P.A., Higham, T.G.F., and Lister, A.M. (2004). Pleistocene to Holocene extinction dynamics in giant deer and woolly mammoth. Nature *431*, 684–689.
11. Bintanja, R., van de Wal, R.S., and Oerlemans, J. (2005). Modelled atmospheric temperatures and global sea levels over the past million years. Nature *437*, 125–128.
12. Berger, G.W., and Pewe, T.L. (2001). Last Interglacial age of the Eva Forest Bed, Central Alaska, from thermoluminescence dating of bracketing loess. Quat. Sci. Rev. *20*, 63–76.
13. Lozhkin, A.V., and Anderson, P.A. (2006). A reconstruction of the climate and vegetation of northeastern Siberia based on lake sediments. Paleontol. Journal *40*, S622–S628.
14. McDowell, P.F., and Edwards, M.E. (2001). Evidence of quaternary climatic variations in a sequence of loess and related deposits at Birch Creek, Alaska: Implications for the stage 5 climatic chronology. Quat. Sci. Rev. *20*, 63–76.
15. Ballard, J.W.O., and Rand, D.M. (2005). The population biology of mitochondrial DNA and its phylogenetic implications. Ann. Rev. Ecol. Evol. Syst. *36*, 621–642.
16. Sher, A., Kuzmina, S., Kuznetsova, T., and Sulerzhitsky, L. (2005). New insights into the Weichselian environment and climate of the Eastern-Siberian Arctic, derived from fossil plants, insects and animals. Quat. Sci. Rev. *24*, 533–569.
17. Drummond, A., Rambaut, A., Shapiro, B., and Pybus, O. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. *22*, 1185–1192.
18. Sjodin, P., Kaj, I., Lascoux, M., and Nordborg, M. (2005). On the meaning and existence of an effective population size. Genetics *169*, 1061–1070.
19. Guthrie, R.D. (2006). New carbon dates link climatic change with human colonization and Pleistocene extinctions. Nature *441*, 207–209.
20. Stuart, A.J., Sulerzhitsky, L., Orlova, L., Kuzmin, Y., and Lister, A. (2002). The latest woolly mammoths (*Mammuthus primigenius*) in Europe and Asia: A review of the current evidence. Quat. Sci. Rev. *21*, 1559–1569.
21. Huelsenbeck, J., and Ronquist, F. (2001). MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics *17*, 754–755.
22. Swofford, D. (2002). PAUP* Phylogenetic Analysis using Parsimony (and Other Methods), Version 4.0b10 ed (Sunderland, Massachusetts: Sinauer Associates).
23. Rambaut, A., and Drummond, A.J. (2004). Tracer, 1.3 ed (Oxford: University of Oxford).
24. Drummond, A.J., Ho, S.Y., Phillips, M.J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. PLoS Biol. *4*, e88.
25. Drummond, A., and Rambaut, A. (2003). BEAST v1.0 (Oxford: University of Oxford).