Nuclear DNA Sequences from Late Pleistocene Megafauna

Alex D. Greenwood,[1] *Cristian Capelli,*† *Göran Possnert,*‡ *and Svante Pääbo**

*Max-Planck-Institute for Evolutionary Anthropology, Leipzig, and Institute of Zoology, University of Munich, Germany; †Istituto di Medicina Legale, Universitá Cattolica del Sacro Cuore, Rome, Italy; and ‡Department of Ion Physics, the Ångström Laboratory, Uppsala, Sweden

We report the retrieval and characterization of multi- and single-copy nuclear DNA sequences from Alaskan and Siberian mammoths (*Mammuthus primigenius*). In addition, a nuclear copy of a mitochondrial gene was recovered. Furthermore, a 13,000-year-old ground sloth and a 33,000-year-old cave bear yielded multicopy nuclear DNA sequences. Thus, multicopy and single-copy genes can be analyzed from Pleistocene faunal remains. The results also show that under some circumstances, nucleotide sequence differences between alleles found within one individual can be distinguished from DNA sequence variation caused by postmortem DNA damage. The nuclear sequences retrieved from the mammoths suggest that mammoths were more similar to Asian elephants than to African elephants.

Introduction

DNA sequences ranging in age from a few hundred to over 50,000 years have been amplified by PCR (e.g., Pääbo 1993; Höss, Pääbo, and Vereshchagin 1994; Höss et al. 1996*a*; Krings et al. 1997). However, most ancient remains contain no endogenous DNA that is retrievable with current techniques (Höss et al. 1996*b*), and in cases in which DNA is present, it is generally degraded to a small average size and extensively modified by hydrolytic, oxidative, and other processes (Pääbo 1989; Lindahl 1993; Poinar et al. 1998). Because of this, DNA sequences that exist in many copies per cell are more easily retrieved than are single-copy sequences, which only exist in two copies per cell in diploid organisms. Consequently, mitochondrial DNA (mtDNA), which exists in several hundred to thousands of copies per cell, has been the main target of study from ancient organisms. Fortunately, mtDNA has many advantages for evolutionary and population genetics (Avise 1991). However, it is clear that the retrieval of nuclear DNA sequences would provide a much larger set of characters for phylogenetic and population genetic analysis and open up the possibility of studying genetic loci involved in determining phenotypic traits in extinct species.

To date, reports of ancient nuclear DNA sequences from mammals have been restricted to the Holocene (the last 10,000 years), except for one report of a β-globin gene sequence from a 12,000-year-old human bone from Morocco (Béraud-Colomb et al. 1995) and histocompatibility gene sequences from a saber-toothed tiger from the La Brea tar pits (Janczewski et al. 1992). In neither case were the DNA sequences extensively characterized. To better assess the feasibility of nuclear DNA retrieval from Pleistocene material, we studied mammoth (*Mammuthus primigenius*) remains found in permafrost deposits. We chose permafrost remains since a low ambient temperature may enhance preservation of biomolecules (Lindahl 1993), and since mammoth mtDNA has been analyzed by several groups (Hagelberg et al. 1994; Höss, Pääbo, and Vereshchagin 1994; Yang, Golenberg, and Shoshani 1996; Derenko, Malyarchuk, and Shields 1997; Ozawa, Hayashi, and Mikhelson 1997; Noro et al. 1998). We show that DNA extracts from mammoths preserved in the permafrost contain fragments of single-copy genes that can be retrieved by PCR. Furthermore, nuclear gene sequences can sometimes be retrieved from some Pleistocene faunal remains found elsewhere.

Materials and Methods

Samples

The molars of two Alaskan mammoths (*Mammuthus primigenius*) (numbers 8460 and 8001) were sampled in the Department of Vertebrate Paleontology of the American Museum of Natural History, New York. The specimens were excavated from a permafrost deposit at Engineer Creek, Alaska. A mammoth mandible and soft tissue from the Novosibirskie Islands and the Indigirka Basin in Siberia were provided by the Mammoth Museum, Yakutia, Russia. A Mylodon sample (BM(NH)M8758) (Höss et al. 1996*a*) was obtained from the Natural History Museum, London, and a cave bear sample (Vi-G3-7012) from Vindija, Croatia, was obtained from the Croatian Academy of Sciences and Arts. Blood from Asian elephants (*Elephas maximus*) was obtained from the Tierpark Hellabrunn, Munich, and a sample from an African elephant (*Loxodonta africana*) was obtained from the Mpala Research Centre, Kenya. Liver from an armadillo (*Cabassous unicinctus*), a three-toed sloth (*Bradypus variegatus*), a two-toed sloth (*Choloepus didactylus*), and a black bear (*Ursus americanus*) were provided by the University of California Museum of Vertebrate Zoology, Berkeley, Calif.

DNA Extraction, PCR, Prevention of Contamination, and Sequencing

Contemporary tissue samples were incubated in 1–2 ml of 10 mM Tris-Cl (pH 7.5), 10 mM EDTA (pH 8.0), 50 mM NaCl, 2% SDS, and 0.6 mg/ml Proteinase

[1] Present address: Department of Mammalogy, American Museum of Natural History, New York, New York.

Key words: ancient DNA, nuclear DNA, mammoth, ground sloth, cave bear, mitochondrial insertions.

Address for correspondence and reprints: Alex D. Greenwood, Department of Mammalogy, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024-5192. E-mail: alexgr@amnh.org.

Mol. Biol. Evol. 16(11):1466–1473. 1999
© 1999 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

Table 1
PCR Primers Used in this Study

Gene	Primer Name	Primer Sequence
Cytochrome *b*	L1 + H1	AAAAAGCTTCCATCCAACATCTCAGCATGATGAAA
		AAACTGCAGCCCCTCAGAATGATATTTGTCCTCA
	L1 + H2	CGTCCAATGTGTGTGTATAGGCAGA
	L2 + H1	CGACAACTACACTCAAACGGAGCATC
16S rDNA	L + H	TTTCGGTTGGGGCGACCTCGGAG
		TTGCGCTGTTATCCCTAGGGTAACT
28S rDNA	L + R	GGTCGTCCGACCTGGGTATA
		TCTAATCATTCGCTTTACCGGAT
vWF	L1 + R1	TTGTGCCGCCCTCAACAGTA
		GTTTACTGCAGTAGAGGTCA
	L2 + R2	CTGCTCCTGACAGCCAGTGAA
		CACCGGGATCACAATAACCT
A2AB	L + R	GGCCGTGGTCATTGGCGTC
		CATGAGGCACCTTGCAGTGCTG
IRBP	L + R	CCATGTCTCCAGCATCCCT
		GCGGTTGTAAATGGTGTCCA
mtDNA nuclear insert	L + R	AGTATATTCGGTCAATGTAGAC
		GAAATAATATCCATGGGTTCAAT

NOTE.—The first two primers in the table are from Kocher et al. (1989). All primers are given in the 5′-to-3′ orientation.

K overnight at 37°C, extracted with phenol and chloroform, and subsequently concentrated to 50 μl using Millipore Ultrafree MC 30,000 NMWL columns. For the ancient samples, 0.2 g of bone powder (collected using a low-speed electric drill) was demineralized at 37°C in 1 ml of 0.5 M EDTA (pH 8.0)/5% sarkosyl with agitation for 2 days. Ten to fifteen microliters of 10 mg/ml Proteinase K was added, and the bone was incubated an additional 2 days at 37°C. After a brief centrifugation, the supernatant was extracted with phenol, phenol/chloroform/isoamyl alcohol, and chloroform/isoamyl alcohol, and silica purification was performed as described (Höss and Pääbo 1993). The first extraction from the Alaskan mammoth was concentrated to approximately 200 μl prior to the silica purification using Centricon-30 microconcentrators (Amicon). For subsequent extractions, the EDTA/sarkosyl solution was replaced after 2 days with a fresh solution containing Proteinase K; the incubation continued for 2 days, and a slightly modified silica protocol was used (Krings et al. 1997). The ancient soft-tissue extraction was performed in the same manner as the bone extraction except that the demineralization step was omitted and the Proteinase K digestion was in 1 ml of 10 mM Tris-HCl (pH 8.0), 2 mM EDTA (pH 8.0), 10 mM NaCl, 1% SDS, and 8 mg/ml DTT. All work was carried out in a laboratory dedicated exclusively to ancient DNA extractions and setting up PCR reactions from ancient DNA. This laboratory is irradiated by ultraviolet light daily, and equipment and working areas are regularly cleaned with bleach. All personnel wear protective clothing and face shields at all times. Mock extractions were performed to monitor contamination. In addition, three extractions from the Alaskan mammoth were performed independently (two by A.D.G. and another by C.C.) to verify results. Amplifications of particular elephant sequences were carried out after the corresponding mammoth work had been completed. The only exceptions were amplifications of the nuclear insertions of a mitochondrial DNA fragment, which have been previously characterized in our laboratory (Greenwood and Pääbo 1999).

Primers used for PCR are given in table 1. For modern samples, 3 μl of extract (or 50 pg of cow DNA or 100 pg of human DNA) was added to 100-μl PCR reactions containing 67 mM Tris-Cl (pH 8.9), 4 mM $MgCl_2$, 16 mM $(NH_4)_2SO_4$, 10 mM β-mercaptoethanol, 10 U *Taq* DNA polymerase, and 25 pmol of each primer. Thirty cycles of 30 s at 95°C, 1 min at 55–60°C, and 1 min at 72°C were followed by 10 min at 72°C. PCR with ancient samples was performed as described (Höss and Pääbo 1993) in 20 or 30 μl reactions. The annealing temperatures used were 55°C for the cytochrome *b,* 57°C for the 16S rDNA, 62°C for the A2AB gene, and 60°C for the remaining amplifications. PCR products were resolved on 3% low-melting-point agarose gels (NuSieve) in a Tris-Acetate buffer and stained with ethidium bromide. Where necessary, reamplifications from gels were performed as described (Höss and Pääbo 1993), and corresponding areas from agarose gels were isolated for the relevant negative controls and reamplified alongside the samples.

PCR products were cloned using the pGEM-T cloning system (Promega). After electroporation into bacteria and ampicillin and blue/white selection, colonies were picked with a sterile pipette tip and added to 30-μl PCR reactions where the M13 forward and reverse primers were used to amplify inserts for 25 cycles using 1.5 U *Taq* DNA polymerase and the buffer described for modern DNA samples. Inserts were sequenced with FITC-labeled M13 universal primers using a PTC-200 thermal cycler (MJ Research). Sequencing reactions were analyzed on an A.L.F. DNA sequencer (Pharmacia). For radioactive sequencing, colonies were inoculated into standard LB ampicillin cultures, and DNA was prepared using the Plasmid Maxi Prep Kit (Qiagen). With the exception of the 28S rDNA clones, the sequences were determined from one strand. However, all sequences were determined from between 4 and

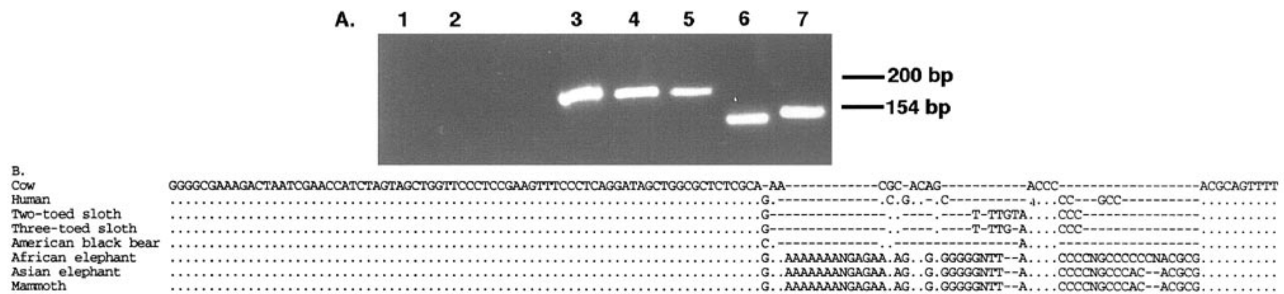

FIG. 1.—*A,* Agarose gel of 28S rDNA PCR products. Templates used were as follows: 1—no template; 2—mock extraction; 3—mammoth extract; 4—Asian elephant; 5—African elephant; 6—cow; 7—human. *B,* Alignment of the 28S rDNA sequences from the mammoth and some mammals. The cow sequence was determined for this study. Dots represent identity to the top sequence and dashes represent gaps. N's represent positions in the Elephantidae homopolymers at which the presence or absence of a base could not be accurately determined.

30 clones, sequence reads were generally short, and for all mammoth nuclear sequences, multiple clones were sequenced from at least two PCR reactions. Radioactive sequencing was done using α-P^{33}-dATP (Amersham) with the T7 Sequenase 2.0 protocol (Amersham). Sequences were resolved on 6% polyacrylamide gels and visualized by autoradiography using XAR film (Kodak).

Results

Dating

The dentine sample from the Alaskan mammoth was dated by accelerator mass spectroscopy at 13,775 ± 145 years, and the bone sample of the Croatian cave bear was dated at 33,335 ± 145 years. The Siberian mammoth and the second Alaskan mammoth were assumed to be of late Pleistocene age. The estimated age of the ground sloth is approximately 13,000 years (Höss et al. 1996*a*).

Mitochondrial Sequences

In order to determine if mammoth DNA was present in the Alaskan mammoth bone extract, a 142-bp fragment (including primers) of the mitochondrial 16S ribosomal DNA (rDNA) and two overlapping fragments of the cytochrome *b* gene (199 and 239 bp, respectively) were amplified. These fragments were readily amplified (not shown). However, an amplification of a 376-bp segment of the cytochrome *b* gene using the two external primers from the previous amplification (table 1) yielded no product. This indicates that the DNA in the extract is degraded to a small size, a phenomenon frequently observed in extracts from ancient samples (Pääbo 1989).

The two cytochrome *b* amplification products were cloned and the inserts of six and eight clones, respectively, were sequenced (not shown). A consensus sequence was generated and compared over a 305-bp segment with three mammoth sequences in the data bank. This sequence differed at one position from those of the other three mammoths, at 12–17 and 9–11 positions from four Asian and four African elephants, respectively, and at 70 and 69 positions from a dugong and a manatee, respectively; the latter are the closest living relatives of mammoths and elephants. The 16S rDNA sequence derived in a similar way from nine clones (not shown) differed by one to six substitutions from four mammoth sequences (Höss, Pääbo, and Vereshchagin 1994) and by four and six substitutions from an Asian elephant and an African elephant, respectively. Dugong or manatee 16S rDNA sequences were not available, but cow 16S rDNA differed by 20 substitutions and a deletion from that of the mammoth. Thus, it was concluded that the predominant mammalian DNA present in this bone is of mammoth origin and that it is abundant enough to be easily amplified by PCR.

Multicopy Nuclear DNA Sequences

In order to test whether nuclear DNA might be preserved in this specimen, we first attempted to amplify the 28S rDNA, which exists in a few hundred copies per haploid genome in vertebrates. When this has previously been attempted for other ancient specimens, it has been noted that the rDNA of fungi can be present in large amounts in ancient samples (Handt et al. 1994). Therefore, primers expected to be specific for vertebrate 28S rDNA (table 1) were constructed by comparing five vertebrate sequences available in GenBank with nonmammalian sequences. When these primers were used in amplifications from the mammoth extract, a product was obtained that was larger than corresponding amplification products from human and cow DNA but similar in size to products from Asian and African elephant DNA (fig. 1*A*). The amplification was repeated for another extract of the same mammoth, and a product of similar size was obtained. However, a smaller amount of a product similar in size to the human sequence was observed, indicating the presence of small amounts of human DNA in the amplifications. A total of 18 clones were sequenced from the first amplification, and 5 were sequenced from the second. The elephant sequences were determined from 10 clones for the Asian and African elephants, respectively. The mammoth and elephant sequences differ from the homologous sequences in the other vertebrates compared in having an expansion segment composed largely of homopolymers. Since the numbers of bases in these homopolymers vary substantially among clones (not shown), they were excluded from further analyses. In the remaining part of the sequences, the mammoth clones differed among themselves by an average of 3.8 substitutions and 1.9 insertions/deletions, whereas the elephant sequences differed by 0.8 and 0 substitutions and 0.36 and 0.31 insertions

for the Asian and African elephants, respectively. Thus, the clones derived from the mammoth show substantially more differences among them than do the elephant clones. This is likely to be the result of two factors. First, the mammoth amplification is likely to have started from much fewer template molecules than the elephant amplifications and thus to have experienced more replications per DNA molecule, increasing the likelihood of polymerase errors during the PCR. Second, the ancient DNA is likely to carry modifications, many of them miscoding (Höss et al. 1996*b*).

When the mammoth consensus sequence was compared with the elephant sequences, it was found to carry one substitution and one deletion of a single base when compared with that of the African elephant and no differences when compared with that of the Asian elephant (fig. 1*B*). In contrast, the human or cow sequences were substantially shorter since they lacked the homopolymer expansions.

Single-Copy Nuclear Sequences

DNA sequences from the 28th exon of the von Willebrand factor gene (vWF) were chosen in order to investigate whether single-copy nuclear genes could be amplified from the mammoth extract since vWF sequences for African and Asian elephants, dugong, and several other mammals are available (Porter, Goodman, and Stanhope 1996). Primers expected to be specific for the elephants were designed to amplify a 96-bp segment (including primers) encompassing three positions that differ between Asian and African elephants. An amplification product was obtained after 40 cycles of amplification, reamplified, cloned, and nine clones were sequenced. The amplification was repeated twice from a second extract, and a total of 21 clones were sequenced (fig. 2). Amplifications from mock extraction, as well as other relevant controls, were negative. The 30 clones differed from each other by 0–9 substitutions. Most of these substitutions occurred in a single clone or very few clones and are therefore likely to be caused by polymerase errors during the amplification, a phenomenon presumably exacerbated by chemical modifications present in the ancient DNA (Höss et al. 1996*b*).

Three substitutions differed from this pattern. Two of these predominated in clones from one amplification but were rare or absent in the other two amplifications (fig. 2, positions marked with asterisks). This is probably due to misincorporation events that occur in an early cycle when amplifications start from few template molecules such that a misincorporated base may come to dominate the population of molecules in the resultant amplification reaction. These types of substitutions are not expected to be consistently reproducible between amplifications (Handt et al. 1996; Krings et al. 1997). The remaining position (fig. 2, position marked "a") differed from this pattern. It appeared as a C in 15 clones, distributed such that 7, 2, and 6 clones carried this base in the three amplifications, whereas it appeared as a T in 14 clones, distributed among 2, 6, and 6 clones in the three amplifications. In addition, one clone carried an A, presumably as a result of polymerase misincorporation. Thus, in total, 15 clones carried a C and 14 clones carried a T at this position. The approximate 1:1 ratio between the two sequences and the fact that both sequences occur in three independent amplifications suggest that two different alleles of the vWF gene were amplified. Consistent with this is the fact that this position carries a T in the dugong sequence and a C in the other mammalian sequences (fig. 2) and that it does not result in any amino acid substitution. Consequently, two consensus sequences, corresponding to the two putative alleles, were constructed from these amplifications. One of these sequences differs by two positions from the African and Asian elephants, whereas the other differs by three positions from both. The dugong differs from the mammoth sequences at 4 and 5 positions, respectively, whereas 24 other mammalian sequences differ at 7–16 positions. Thus, the sequences determined from the mammoth are most similar (but not identical) to the proboscidean species to which the mammoth is most closely related.

Primers were designed to amplify a second segment located farther downstream in the 28th exon of the vWF. This segment is 99 bp long (including primers) and encompasses four positions that differ between Asian and African elephants. Three amplifications from two mammoth extracts yielded PCR products, which were cloned, and a total of 24 clones were sequenced (fig. 2). Amplifications from the mock extraction and relevant controls were negative. The clones differ from each other at 0–10 positions. Three positions showed a pattern interpreted to reflect misincorporations in early cycles of the PCR. One position is of particular interest, since it carries an A in all 10 clones from one amplification but in only one of 14 clones from the other two amplifications. In this case, sequencing of more clones from several independent amplifications might be necessary to elucidate whether this represents an allelic difference or a polymerase error that affected most clones in the first amplification and very few clones in the second amplification. Irrespective of this, a consensus sequence, which may thus represent one or both of the alleles present in this mammoth extract, was constructed for all 24 clones. It was found to differ by two substitutions from those of each of the elephants and the dugong, but by 5–17 differences from 25 mammalian sequences in the database. Again, the sequence is similar but not identical to those of the living species most closely related to the mammoth.

Primers were designed to amplify a 98-bp segment of the alpha-2B adrenergic receptor (A2AB) gene (Stanhope et al. 1998), and three amplifications were performed from two extracts. The PCR products were cloned, and a total of 26 clones were sequenced. All relevant controls were negative. Five positions carried one base in many clones of a given PCR product and a different base in the clones from the other two PCR products (fig. 2). A consensus sequence was reconstructed based on the assumption that these positions represent misincorporations. As the African elephant sequence was not available, it was amplified and sequenced subsequent to the mammoth experiments. The

```
vWF gene, 28th exon, 5' fragment
                          a          *                    *
Asian elephant         GCCCC-AGTC-AGCGTCACCACGCCTTTTGTGGAGGACATCTCAGAGCCGCCCCTGSA
Cow                    .G...-G...-.T.CC...A..C..A.A............C.G...........C.
Human                  .....-G..G-...CC......T.TG.A...............G..A.....GT..C.
Dugong                 .....-....-..T.C.........A.A............C......T.......C.
African elephant       .....-....-..............A.A...................T.......C.
Mammoth,
consensus sequence
allele 1:              .....-....-..............G.A...........................C.
allele 2:              .....-....-..T...........G.A...........................C.

Mammoth,               .....-....-..............G.A.A..................A......C.
clones:1st extract,    .....-....-..........T...G.A...........................C.
1st PCR                .....-....-..T...........G.A...........................C.
                       .....-....-......N.......G.A.A..................A......C.
                       .....-...T-..............G.A.AA.................A......C.
                       .....-....-..T...........G.A......A....................C.
                       .....-....-..............G.A.A..................A......C.
                       .....-....-..............G.A.A..................A......C.
                       .....-....-..............G.A.A..................A......C.

Mammoth,               .....-....-..T.......T...G.A...............G...........C.
clones:2nd extract,    .....-....-..T..........TG.A...........................C.
1st PCR                .....C...TN..T.....T.T...G.A...........................C.
                       .....-....-..............G.A...........C...............C.
                       .....-....-..T...........A.A...........................C.
                       .....-....-..............G.A...........................C.
                       .....-....-..T...........G.A...........................C.
                       .....-....-..A...........G.A....A......................C.
                       ..N..-....-..T...........G.A...........................C.

Mammoth,               ...T.-....-..............G.A...........................C.
clones:2nd extract,    .....-....-..............G.A...........................C.
2nd PCR                .....-....-..T...........G.A......................T.T..C.
                       .....-....-..............G.A.A.AA......................C.
                       .....-....-..T...........G.A...........................C.
                       .....-....-..T....N......G.A.N.................T.......C.
                       .....-....-..T...........A.A....................A......C.
                       .....-....-..T...........G.A.............T........T.T..T.
                       .N...C....-..............G.A...........................C.
                       .....-....-..............G.A..........G................C.
                       .....-....-.......T..T..TG.A...................T.......C.
                       .....-....-..T....TT....TG.A...........................C.
```

```
vWF gene, 28th exon, 3' fragment
 *       *         *
GAGCCCCC-ACGGATGGCCCGGAAATTGG-TCCGATATGTGCAGGGCCTGAAGAAGAAGA
..A.....-CA..C.........C....-....C..C..C...................
.......A-.......T.......C..T.-....C..C..C...................
........-...............C....-....C.........................
........-....T..........C....-C.........C...................

........-...............C....-C.............................

........-.T.........A...C....-C.............................
........-.T.........A...C....-C.............................
.......N-...........A...C....-C.............................
.......--......A....A...C....-C.............................
.......--...........A...C....-C......C......................
........N.T.........A...C....-C.............................
.......--.T.........A...C....-C.............................
........-...........A...C....-C.............................
.......--.....C.....A..GC....-C.............................
......---........... A...C....-C..................T..........

..A.....-...............C....-C................A..........A.
........-...............C....-C......A......................
..A.....-...............C....-C..................G.........
..A.....-......A........C....-C.............................
........-...........A...C.....C.....................A.......
....T...-...............CC...-C.T.............T............G

........-.T.............C....-C.............................
........-.....A.........C....-C.............................
........-...............C....NC.............................
T.......-.........TT....C....-C.............................
........-...............C....-C.......................A....
........-...............T....-C...G.........................
........-...............C....-C.............................
...T....-...............T....-.TT........T..................
```

```
A2AB gene
                        * *                        *       *        *
Asian elephant         TTTGTGCTCTGCTGGTTCCCGTTCTTCTTCAGCTACAGTCTGGGTGCCATTTGCCCG
Cow                    ..C................T.................C..........C.....
Human                  ...................C.................C.....C....C.....
Dugong                 ..C................T.................C..........C....A
African elephant       .......................................................
Mammoth,
consensus sequence     .......................................................

Mammoth,               .......................................................
clones:2nd extract,    ....................................AA.............
1st PCR                .......................................................
                       ........................A..............................
                       ..............................A........................
                       .............A.........................................
                       ........T..............................................
                       .......................................................

Mammoth,               ..................................................-.TT.
clones:3rd extract,    .......................................................
1st PCR                ..................................................A....
                       ..........................................A........A....
                       ............................................A..........
                       ..........................................A......C.A....
                       .......................................................
                       ..........................................A........A....
                       .......................................................

Mammoth,               ......T.T.........................T....................
clones:3rd extract,    ......T.T.........................T....................
2nd PCR                ......T.T.........................T...................A
                       ......T.T.........................N....................
                       ......T.T.........................T....................
                       ......T.T............A............T....................
                       ......T.T.........................T....................
                       ......T.T.........................T....................
                       ....................A..................................
```

```
IRBP gene

TACCTTATTTCCTACCTGCACCCGGGCGGCACCGTGCTGCACG
..TG.C..C.............A..GA....A..CTC.....
...A.C..C.............A..GAA....A.C.......
...A.C..C........A.....C..A..AG.A........T.
.....C....................................

..........................................

.....................................T...
.....................................T...
..........................................
..........................................
.................T.TTT......T.............
................................A.A.......
...........................A..............
............................N........A....
..........................................

...........-......T..T....................
..........................................
..........................................
..........................................
................................A.........
.......................A........A...A....
..T.................T.....................
..........................................
..........................................
..........................................
..........................................
..........................................
```

```
mtDNA insertion

Asian elephant         AGTCATGTAATTCACGAATCCCC-TAACACTT
African elephant       .....C.................-........
Mammoth,
consensus sequence     .......................-........

Mammoth,               .......................-........
clones:1st extract,    T......................-........
1st PCR                C......................-........
                       ...........C...........-........
                       .......................-........
                       ........G..........N...C........
                       .......................-........
                       .......................-........

Mammoth,               C...............G......-........
clones:1st extract,    .......................-........
2nd PCR                .......................-........
                       .......................-........
                       .......................-........
                       .......................-........
                       .......................-........
                       .......................-........
                       .......................-........
                       ...T.....-.............-........
```

FIG. 2.—DNA sequences for four single-copy nuclear sequences and one mitochondrial insertion from the mammoth. In each case, the DNA sequences of the clones used to reconstruct the mammoth sequence are given below. Above, the mammoth sequences are compared with relevant mammalian sequences. Dots represent identity to the Asian elephant sequence, dashes represent gaps, S represents an ambiguity in the

mammoth sequence was found to be identical to both the African and the Asian elephants and to differ at five positions from the dugong and at four to eight positions from 13 other mammalian taxa in the database.

Primers for an 82-bp fragment of the interphotoreceptor retinoid binding protein (IRBP) gene were designed. The product was amplified twice from one extract and directly cloned without reamplification (fig. 2). Amplifications from mock extractions, as well as other relevant controls, were negative. Since no consistent differences between clones from the two amplification products were observed, a third PCR was not performed. Subsequent to the mammoth amplifications, the Asian elephant sequence, which was not available in the database, was amplified and sequenced. The mammoth and the Asian elephant sequences were identical, whereas they differed at 1 position from the African elephant, at 10 positions from the dugong, and at 5–14 positions from 29 other mammalian taxa in the database.

Nuclear Insertions of Mitochondrial DNA

Insertions of mitochondrial DNA into the nuclear genome are found in most species and can be mistaken for bona fide mitochondrial sequences (Zischler et al. 1995; Zhang and Hewitt 1996). It has recently been shown that elephants carry a large number of nuclear insertions derived from hypervariable region I of the mtDNA (Greenwood and Pääbo 1999). Primers were designed to amplify a 76-bp segment (including primers) of one such insertion that exists in Asian and African elephants. The copy number of this nuclear sequence is not known, but it is probably lower than that of 28S rDNA. Two amplifications were performed, and a total of 18 clones were sequenced (fig. 2). Amplifications from mock extraction and all other relevant controls were negative. The consensus sequence was identical to the Asian elephant insertion sequence and carried one difference from the same insertion sequence in an African elephant.

Sequence Comparisons

When all sequences from single-copy genes were concatenated from the mammoth, the two elephants, and the dugong, 31 positions were found to be variable. The mammoth carries 4–5 substitutions to the Asian elephant, 5–6 substitutions to the African elephant, and 21–22 substitutions to the dugong. However, since only three positions were informative in a parsimony sense, further phylogenetic analyses were not performed. It is, however, of interest that the mammoth shows fewer differences from the Asian elephant than from the African elephant. This remains true when the 28S rDNA and nuclear mtDNA insertion sequences are considered, since each of these carry one substitutional difference from the African elephant and none from the Asian elephant. The determination of further nuclear sequences will be necessary to clarify the phylogenetic relationship of mammoths and the extant elephants.

Additional Samples

Further preliminary experiments indicate that the success with this mammoth is not an isolated occurrence. Extracts prepared from the dentine of another Alaskan mammoth and from the mandible of a Siberian mammoth yielded amplification products of cytochrome *b,* 28S rDNA, and the vWF gene (not shown). Furthermore, among eight extracts prepared from Siberian mammoths that were tested for cytochrome *b* and 28S rDNA amplifications, two failed to yield amplification products from either DNA sequence, and three were positive for cytochrome *b* but not for 28S rDNA, whereas the remaining three yielded amplification products for both the mitochondrial and the nuclear sequences. The latter products were all of a size expected for proboscideans (unpublished observation).

In order to elucidate whether nuclear DNA sequences might also be preserved in nonpermafrost remains, DNA was extracted from a bone sample of a 13,000-year-old ground sloth found in a cave in southern Chile. MtDNA has previously been retrieved from this animal in two different laboratories (Höss et al. 1996*a*; Taylor 1996). Using the 28S rDNA primers, a band was obtained that was shorter than the elephant and mammoth amplification products. When the sequence was reconstructed from nine clones (not shown) and compared with the homologous sequences determined from a two-toed sloth and a three-toed sloth, it was found to lack the homopolymer expansions. It differed at 4 and 5 positions from those of the three- and two-toed sloths, respectively, and at 11 or more positions from other mammals. Thus, the Mylodon sequence was most similar to those of the extant sloths. Attempts to amplify a vWF gene sequence from the ground sloth extract failed.

The 28S rDNA primers were furthermore used in amplifications from the extract of a 33,000-year-old bone from a cave bear (*Ursus spelaeus*) from Croatia. Previously, a mtDNA sequence has been described from a French cave bear specimen (Hanni et al. 1994). A faint band was visualized from the amplification that differed in size from those of elephants, mammoths, cows, sloths, and humans. The sequence was reconstructed from the cloned amplification product and was found to be identical to the homologous sequence amplified subsequently from an American black bear (*Ursus americanus*). As the mitochondrial and 28S rDNA amplifications were weak, amplification of single-copy nuclear sequences was not attempted.

Discussion

The results presented show that nuclear DNA sequences that exist in several copies per haploid genome,

←

published Asian elephant sequence, and N's represent bases that could not be accurately determined. Asterisks indicates bases that appear in several clones in a given PCR amplification but differ from most clones from other PCR reactions, whereas "a" denotes a position that carries two different bases in multiple clones from three PCR reactions.

such as 28S rDNA, as well as sequences that exist as single copies, such as the vWF, A2AB, and IRBP genes, can be amplified by PCR from a mammoth sample. Nonpermafrost samples also yielded nuclear sequences. Thus, multicopy nuclear DNA seems to be retrievable not only from permafrost remains, but also from well-preserved bones found in other contexts. However, it is notable that soft-tissue remains from a Siberian mammoth failed to yield any ancient sequences. It may be that bone tends to preserve DNA better than soft tissues, as has been previously reported (Cooper et al. 1992).

The primers used for 28S rDNA amplifications may be useful for screening extracts of mammalian remains for amplifiable nuclear DNA, since their amplification products differ in size among species, for example, mammoths as compared with several other mammals. Thus, a simple electrophoretic analysis of the amplification product can be used to distinguish amplification products derived from endogenous DNA from some of those derived from contaminants (fig. 1*A*). In this respect, humans and cows are of particular interest, since the former often contaminate palaeontological remains (Krings et al. 1997) and the latter may derive from bovine serum albumin, which is often added to the PCR in order to decrease the effects of inhibitory factors present in extracts of ancient tissues (Pääbo 1990). Thus, the primers described could be used to screen ancient samples for potential endogenous nuclear DNA and only products of a size compatible with the species in question submitted to sequencing.

When a putative nuclear DNA sequence has been identified, the results presented demonstrate that it is of crucial importance that the DNA sequences of multiple clones derived from several PCR reactions are determined in order to distinguish allelic variation from PCR errors. The latter may be derived both from nucleotide misincorporations and from ''jumping PCR'' (Pääbo, Irwin, and Wilson 1990) taking place between alleles or members of a gene family or other repeated DNA sequences.

The amplification of a nuclear insert of a mitochondrial sequence from the mammoth indicates a potential problem for studies of mtDNA from ancient remains. Since single-copy sequences can be obtained from some ancient samples, the possibility exists that when primers designed to amplify mitochondrial sequences are used in amplifications from such samples, some of the sequences amplified could be of nuclear, rather than mitochondrial, origin. This possibility is enhanced by the fact that some mitochondrial insertions are present in multiple copies in the nuclear genome (Lopez et al. 1994). It is therefore important that precautions, for example, the use of several primer pairs for generating overlapping pieces of a template sequence (Handt et al. 1996), are used when mtDNA from ancient remains is studied.

Conclusions

A plethora of faunal remains exist in permafrost deposits and other cold environments. The fact that such remains can yield not only mtDNA but also single-copy nuclear DNA sequences in a substantial amount of cases opens up the possibilities of using nuclear loci in phylogenetic and population genetic studies and of studying genes determining phenotypic traits. However, the sequencing of multiple clones from several PCR products derived from each nuclear DNA sequence under study is necessary in order to distinguish allelic variation from *Taq* DNA polymerase errors.

GenBank Accession Numbers

Mammoth (*M. primigenius*): cytb, AF154864; 16S rDNA, AF154865; 28S rDNA, AF154872; vWF 5′-fragment allele 1, AF154873; vWF 5′-fragment allele 2, AF154874; vWF 3′ fragment, AF154975; IRBP, AF155042; A2AB, AF154876; mtDNA insert, AF155040. Asian elephant (*E. maximus*): 28S rDNA, AF154870; IRBP, AF155043. African elephant (*L. africana*): 28S rDNA, AF154871; A2AB, AF154877; mtDNA insert, AF155041. Cow (*Bos* sp.): 28S rDNA, AF154866. Two-toed sloth (*C. didactylus*): 28S rDNA, AF154867. Three-toed sloth (*B. variagatus*): 28S rDNA, AF154868. Black bear (*U. americanus*): 28S rDNA, AF154869.

Acknowledgments

The authors thank N. Georgiadis and J. Hektor for elephant samples; R. Tedford, G. Boeskerov, and P. Lazarev for mammoth samples; K. Großschmidt and M. Paunovic for the cave bear sample; G. Feldmaier-Fuchs for technical assistance; R. Evander, C. Flemming, H. Kaessmann, M. Krings, R. MacPhee, A. Reichert, M. Stoneking, K. Strimmer, L. Vigilant, and J. Wilson for assistance and helpful discussions; the Deutsche Forschungsgemeinschaft for financial support; the Irene Diamond Fund; and the National Science Foundation for an Alfred P. Sloan Postdoctoral Fellowship in Molecular Evolution to A.D.G.

LITERATURE CITED

AVISE, J. C. 1991. Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. Annu. Rev. Genet. **25**:45–69.

BÉRAUD-COLOMB, E., R. ROUBIN, J. MARTIN, N. MAROC, A. GARDIESEN, G. TRABUCHET, and M. GOOSSENS. 1995. Human beta-globin gene polymorphisms characterized in DNA extracted from ancient bones 12,000 years old. Am. J. Hum. Genet. **57**:1267–1274.

COOPER, A., C. MOURER-CHAUVIRE, G. K. CHAMBERS, A. VON HAESELER, A. C. WILSON, and S. PÄÄBO. 1992. Independent origins of New Zealand moas and kiwis. Proc. Natl. Acad. Sci. USA **89**:8741–8744.

DERENKO, M., B. MALYARCHUK, and G. F. SHIELDS. 1997. Mitochondrial cytochrome b sequence from a 33,000 year-old woolly mammoth (*Mammuthus primigenius*). Ancient Biomolecules **1**:149–153.

GREENWOOD, A. D., and S. PÄÄBO. 1999. Nuclear insertion sequences of mitochondrial DNA predominate in hair but not in blood of elephants. Mol. Ecol. **8**:133–137.

HAGELBERG, E., M. G. THOMAS, C. E. J. COOK, A. V. SHER, G. F. BARYSHNIKOV, and A. M. LISTER. 1994. DNA from ancient mammoth bones. Nature **370**:333–334.

HANDT, O., M. HÖSS, M. KRINGS, and S. PÄÄBO. 1994. Ancient DNA: methodological challenges. Experientia **50**: 524–529.

HANDT, O., M. KRINGS, R. H. WARD, and S. PÄÄBO. 1996. The retrieval of ancient human DNA sequences. Am. J. Hum. Genet. **59**:368–376.

HANNI, C., V. LAUDET, D. STEHELIN, and P. TABERLET. 1994. Tracking the origins of the cave bear (*Ursus spelaeus*) by mitochondrial DNA sequencing. Proc. Natl. Acad. Sci. USA **91**:12336–12340.

HÖSS, M., A. DILLING, A. CURRANT, and S. PÄÄBO. 1996*a.* Molecular phylogeny of the extinct ground sloth *Mylodon darwinii.* Proc. Natl. Acad. Sci. USA **93**:181–185.

HÖSS, M., P. JARUGA, T. H. ZASTAWNY, M. DIZDAROGLU, and S. PÄÄBO. 1996*b.* DNA damage and DNA sequence retrieval from ancient tissues. Nucleic Acids Res. **24**:1304–1307.

HÖSS, M., and S. PÄÄBO. 1993. DNA extraction from Pleistocene bones by a silica-based purification method. Nucleic Acids Res. **21**:3913–3914.

HÖSS, M., S. PÄÄBO, and N. K. VERESHCHAGIN. 1994. Mammoth DNA sequences. Nature **370**:333.

JANCZEWSKI, D. N., N. YUHKI, D. A. GILBERT, G. T. JEFFERSON, and S. J. O'BRIEN. 1992. Molecular phylogenetic inference from saber-tooth cat fossils of Rancho La Brea. Proc. Natl. Acad. Sci. USA **89**:9769–9773.

KOCHER, T. D., W. K. THOMAS, A. MEYER, S. V. EDWARD, S. PÄÄBO, F. X. VILLABLANCA, and A. C. WILSON. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA **86**:6196–6200.

KRINGS, M., A. STONE, R. W. SCHMITZ, H. KRAINITZKI, M. STONEKING, and S. PÄÄBO. 1997. Neanderthal DNA sequences and the origin of modern humans. Cell **90**:19–30.

LINDAHL, T. 1993. Instability and decay of the primary structure of DNA. Nature **362**:709–715.

LOPEZ, J. V., N. YUHKI, R. MASUDA, W. MODI, and S. J. O'BRIEN. 1994. *Numt,* a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. J. Mol. Evol. **39**:174–190.

NORO, M., R. MASUDA, I. A. DUBROVO, M. C. YOSHIDA, and M. KATO. 1998. Molecular phylogenetic inference of the woolly mammoth *Mammuthus primigenius,* based on complete sequences of the mitochondrial cytochrome b and 12S ribosomal RNA genes. J. Mol. Evol. **46**:314–326.

OZAWA, T., S. HAYASHI, and V. M. MIKHELSON. 1997. Phylogenetic position of mammoth and steller's sea cow within tethytheria demonstrated by mitochondrial DNA sequences. J. Mol. Evol. **44**:406–413.

PÄÄBO, S. 1989. Ancient DNA: extraction, characterization, molecular cloning and enzymatic amplification. Proc. Natl. Acad. Sci. USA **86**:1939–1943.

———. 1990. Amplifying ancient DNA. Pp. 159–166 *in* M. INNIS, D. GELFAND, J. SNINSKY, and T. WHITE, eds. PCR protocols: a guide to methods and applications. Academic Press, Orlando, Fla.

———. 1993. Ancient DNA. Sci. Am. **269**:86–92.

PÄÄBO, S., D. M. IRWIN, and A. C. WILSON. 1990. DNA damage promotes jumping between templates during enzymatic amplification. J. Biol. Chem. **265**:4718–4721.

POINAR, H., M. HOFREITER, W. G. SPAULDING, P. S. MARTIN, B. A. STANKIEWICZ, H. BLAND, R. P. EVERSHED, G. POSSNERT, and S. PÄÄBO. 1998. Molecular coproscopy: dung and diet of the extinct ground sloth *Nothrotheriops shastensis.* Science **281**:402–406.

PORTER, C. A., M. GOODMAN, and M. J. STANHOPE. 1996. Evidence on mammalian phylogeny from sequences of exon 28 of the von Willebrand factor gene. Mol. Phylogenet. Evol. **5**:89–101.

STANHOPE, M. J., V. G. WADDELL, O. MADSEN, W. DE JONG, S. B. HEDGES, G. C. CLEVEN, D. KAO, and M. S. SPRINGER. 1998. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. Proc. Natl. Acad. Sci. USA **95**:9967–9972.

TAYLOR, P. 1996. Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. Mol. Biol. Evol. **13**:2839.

YANG, H., E. M. GOLENBERG, and J. SHOSHANI. 1996. Phylogenetic resolution within the Elephantidae using fossil DNA sequences from the American mastodon (*Mammut americanum*). Proc. Natl. Acad. Sci. USA **93**:1190–1194.

ZHANG, D., and G. M. HEWITT. 1996. Nuclear integrations: challenges for mitochondrial DNA markers. Trends Ecol. Evol. **11**:247–251.

ZISCHLER, H., M. HÖSS, O. HANDT, A. VON HAESELER, A. C. VAN DER KUYL, J. GOUDSMIT, and S. PÄÄBO. 1995. Detecting dinosaur DNA. Science **268**:1192–1193.

CHARLES AQUADRO, reviewing editor

Accepted July 7, 1999