



Supporting Online Material for

Comment on “Whole-Genome Shotgun Sequencing of Mitochondria from Ancient Hair Shafts”

Regis Debruyne,* Carsten Schwarz, Hendrik Poinar

*To whom correspondence should be addressed. E-mail: debruyne@mcmaster.ca

Published 7 November 2008, *Science* **322**, 857a (2008)
DOI: 10.1126/science.1158417

This PDF file includes:

Materials and Methods
SOM Text
Figs. S1 and S2
Tables S1 to S8
References

Supplementary Online Material

Materials and Methods

Samples

Eight samples from four different mammoths were used and are listed in Table S2. Two of those mammoths were extracted from bone only: 2003/384 also known as the Fishhook mammoth of which hair was extracted by Gilbert *et al.* (1), and 2005/915, formerly used by Poinar *et al.* (2) in their metagenomic study and used for comparison with hair samples by Gilbert *et al.* (1).

The two other mammoths were extracted from both hair and bone: the so-called Jarkov mammoth, discovered in 1997 and preserved since then in frozen conditions (its hair was used by Gilbert *et al.* (1)), and the Lyakhov mammoth, discovered in 1902 on the Main Lyakhov Island (Siberia) and preserved at room temperature since then, having been offered to the French Museum of Natural History (Paris) in 1912 where it is still on exhibit.

DNA extraction from hair/root

Individual mammoth hair weighed between 0.1 mg (underfur hair) and 12 mg (coat hair). Hair samples were prepared according to Gilbert *et al.*'s protocol (1): roots were excised (although not discarded but pooled together and extracted separately from the hair shaft), and hair shafts were soaked with 5% bleach before being rinsed in nanopure water (Barnstead). The only adjustment compared to the original protocol consisted in an ultrafiltration step (with concentration to a final volume of 100 μ l eluted in 0.1X TE, pH 8.0) on YM-30 Microcon filtered columns (Millipore) rather than on the YM-10 Amicon (Millipore) filtered columns. Remarkable improvement in the removal of PCR-inhibitors justified that choice: in PCR, hair shaft extracts could therefore be used 10 times more concentrated than when ultrafiltered on YM-10 Amicon without sacrificing DNA yield as determined by qPCR (see inhibition test below for details).

DNA extraction from bone

Bone samples were extracted according to the protocol published by Poinar *et al.* (2). Samples (100-110 mg of dense bone) were first demineralized in 1.5 ml of 0.5M EDTA (pH 8.0) overnight under slight agitation at room temperature. After removal of the supernatant, the bone pellet was digested in 1.5 ml of digestion buffer at 55°C until completion of the digestion (between 3 and 5 hours). Both supernatant and pellet digests were subsequently extracted with 500 μ l of phenol/chloroform/isoamyl alcohol (25:24:1, pH 8.0) and 500 μ l of chloroform before filtration and concentration to a final volume of 100 μ l in 0.1X TE (pH 8.0) by ultrafiltration using a Microcon YM-30 column (Millipore).

Measure of total DNA concentration of the extracts

The total yield of DNA per extract (Table S3) was quantitated for each extract using a PicoGreen assay (Invitrogen) at a dilution factor of 1:50; 1 μ l of extract

was diluted in 24 ul of 1X TE prior to being mixed with 25 ul of PicoGreen dye solution, as per manufacturer's recommendations. After 5 minutes of incubation at room temperature, the DNA concentration was measured on a fluorometer (TBS-380, Turner Biosystems) calibrated with a 10 ng.ul⁻¹ reference DNA solution.

Quantitative PCR methodology

Quantitative PCR using SYBR green as a reporter dye was used to quantitate mammoth DNA. Primers are listed in Table S4.

Quantitative PCR reactions were performed in 20 ul, using 4 ul of template DNA and 1 U of AmpliTaq Gold polymerase (Applied Biosystems), 1X of PCR buffer, 25 mM of MgCl₂, 0.6 ug.ul⁻¹ of Bovine Serum Albumin, 300 nM of both forward and reverse primers (Integrated DNA Technologies), 250 uM of dNTPs (GE Healthcare), 30 nM of ROX reference dye, 0.167X of SYBR Green dye. Assays were performed on the 7900 HT real-time PCR system (Applied Biosystems) using the SDS software (ver2.2.1) with automatic calculation of the threshold and baseline parameters.

Cycling conditions were as follows: 7 minutes of initial denaturation at 95°C, 45 cycles of 30 second denaturation at 95°C, 30 second annealing (at 60°C for all assays) and 90 seconds extension at 72°C (at the end of which fluorescence values were collected at each cycle). A dissociation curve followed the PCR amplification. All PCR conditions had been previously tested on 20 ng of human whole genomic DNA and yielded negative results.

Standard curves were generated using a cloned PCR product amplified from modern elephant DNA spanning all primer binding sites (quantitated using the above mentioned PicoGreen assay). Standards, samples, extraction blanks and PCR blanks were all performed in duplicates (see Table S5 for original Ct values).

PCR Inhibition test

Before the actual qPCR analyses were performed, serial dilutions of the extracts (from straight up to 1:5000) were tested on the shortest (84 bp) mitochondrial assay using standard cycling conditions and reactions in order to address the levels of PCR-inhibition in all of them. The final dilution used was selected to maximize the total yield and minimize the inhibition level and maximize quantitation accuracy (Table S6).

Quantitation rescaling

The final quantitation presented in the Table 1 of the main text were rescaled from the qPCR output (Table S5) to provide a concentration per one mg of starting material (Table S7):

$$\text{Rescaled copy\#} = ((\text{Raw copy\#} * (100 / 4)) * \text{dilution factor}) / \text{amount of material}$$

Modelling of the mitochondrial DNA degradation

Deagle et al. (3) have proposed a statistical model which describes the post-mortem degradation of DNA molecules under the assumption that this is a random process, based on the comparative analysis of the copy number for

different DNA amplicon sizes. In such a model, the distribution of the fragment size of amplifiable DNA molecules can be fitted with a Poisson distribution of parameter λ , where λ is an estimate of the frequency of non-bypassable damage. In order to calculate λ , a linear regression of natural log-transformed copy numbers against fragment length is performed. If the coefficient of correlation of that regression is satisfactory (in our case all R^2 values are greater than 0.97), then λ can be estimated by the slope of the regression line. The inverse of λ can also be used to describe the data: it corresponds to the mean amplifiable fragment size in the sample. The Y-intercept of the regression line ($\ln(N)$) can be used to derive N , a measure for the quantity of DNA (compared to λ being a qualitative parameter). Table S1, showing the results, is discussed in the main text. Based on the obtained data, this model predicts a maximum amplicon size of 960 bp for the Jarkov's bone but only 670 bp for its hair (provided primers and PCR conditions are optimized to a sensitivity down to one single copy).

Supporting text

After a thorough examination of the data presented by Gilbert et al. (1), several unanswered questions remain:

1. There is tremendous variation in the starting amounts of hair shafts used between samples (ranging from 0.2 to 5.2 g), and it is impossible to locate what specific amounts for which samples and how these variations were they accounted for in the comparisons?
2. How much DNA did each library contain?
3. What was the total number of reads per run for each library and how many emulsion PCRs was this?
4. How many identified endogenous mammoth DNA reads versus exogenous DNA reads were recovered? What is the impact of the nebulization on the bone DNA library and on the percent endogenous?
5. What was the ratio of nuclear to mitochondrial mammoth DNA?
6. How precisely were the C>T rate calculations performed?

The lack of these specifics (most of them being essential parameters in the 454 procedure), in addition to the points raised in the main text, precludes the significance of the cross comparisons performed in Gilbert *et al.*'s paper. For example, although the endogenous mtDNA content of hair shaft appears to be relatively higher than in bone (up to 2% of the sequences in contrast to 0.08), this enrichment might result from the conjunction of (i) the relative richness of mtDNA copies in hair tissues compared to bone and (ii) the enrichment of the bone library with contaminant during the nebulization process.

Supporting figures

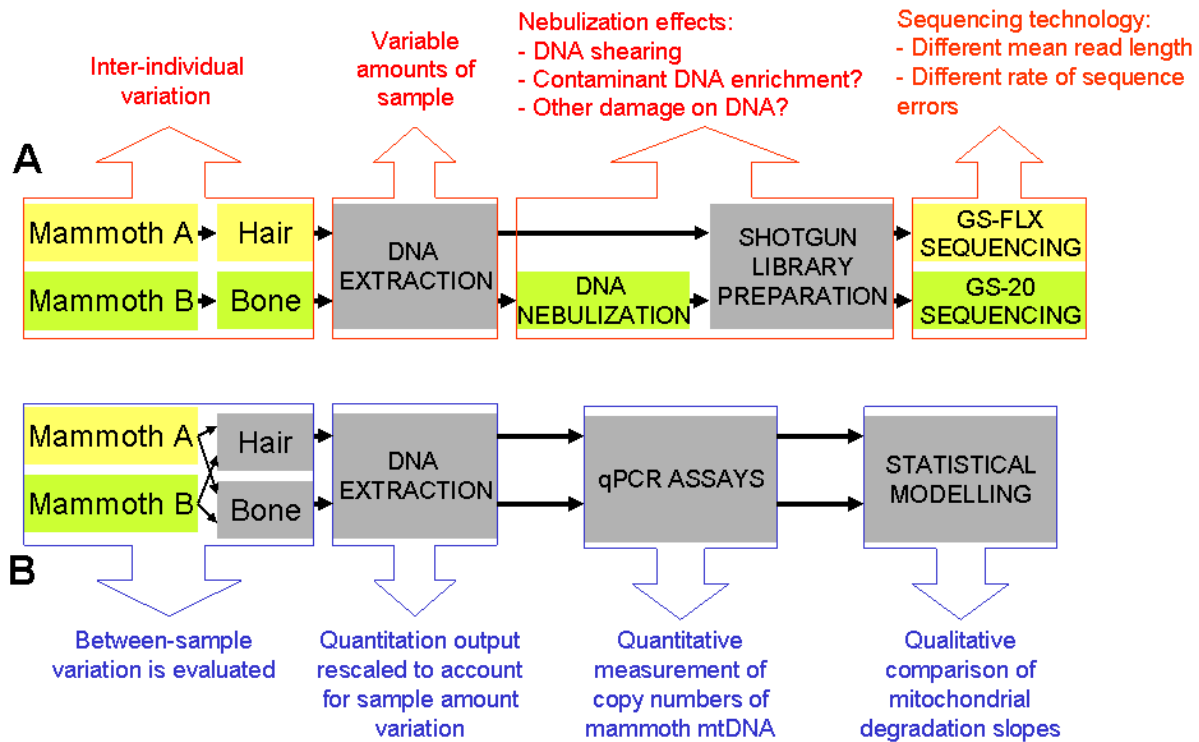


Figure S1

Diagrammatic representation of Gilbert *et al.*'s (A) and the present (B) experimental designs with free (red) and controlled (blue) parameters displayed.

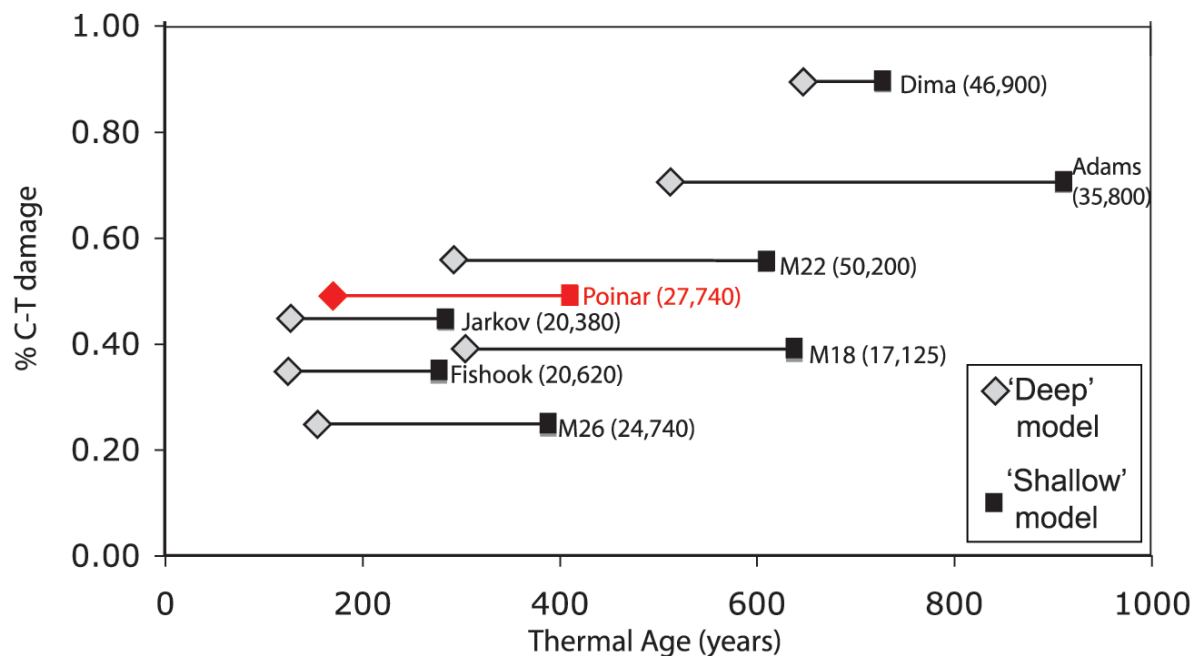


Figure S2

Comparison of thermal age estimates against percentage C-T damage using the estimate for the “Poinar” mammoth (sample 2005/915, here in red) generated in that study (see main text). This figure is modified after Gilbert *et al.*’s figure 2 (p.1929) and thermal ages estimates are all derived from their two alternative models. Contrary to what Gilbert *et al.* (1) state, the level of damage is no longer lower in HS libraries than in bone once our new rate is implemented for the so-called “Poinar” mammoth.

Supporting tables

Table 1: Quantitative and qualitative results for all samples

All results shown are scaled to copies per mg of sample material. Quantitative PCR results (average copy number [c#] of duplicates, and associated standard deviation) are provided by fragment length assayed. Parameters N and λ refer to Deagle *et al.*'s regression model of degradation of which the coefficient of correlation (R^2) is given (3). *Regression lines were generated out of only two fragment sizes (for the shortest assays), which makes calculation of R^2 irrelevant.

Specimen	Sample type	Mitochondrial DNA					Degradation modelling			
		84 bp c#.mg ⁻¹	151 bp c#.mg ⁻¹	279 bp c#.mg ⁻¹	490 bp c#.mg ⁻¹	677 bp c#.mg ⁻¹	N c#.mg ⁻¹	λ bp ⁻¹	1/ λ bp	R ²
Lyakhov	Root	892.4 ±26.9	104.7 ±55.1	0.0 ±0.0	0.0 ±0.0	0.0 ±0.0	13643	0.0325	31	1.00*
	Shaft	114.3 ±21.3	14.7 ±20.9	0.0 ±0.0	0.0 ±0.0	0.0 ±0.0	1547	0.0310	32	1.00*
	Bone	6277.0 ±1817.7	758.8 ±124.7	0.0 ±0.0	0.0 ±0.0	0.0 ±0.0	92411	0.0320	31	1.00*
Jarkov	Root	166804.8 ±94902.3	33441.4 ±16863.9	936.1 ±453.2	0.0 ±0.0	0.0 ±0.0	1677810	0.0267	37	0.999
	Shaft	382901.6 ±22049.0	64526.1 ±6449.0	2810.0 ±316.0	58.0 ±82.1	0.0 ±0.0	1746283	0.0215	47	0.993
	Bone	24119.6 ±3902.3	7392.3 ±851.4	3047.6 ±1225.1	196.7 ±26.4	6.9 ±5.5	55437	0.0114	88	0.986
Fishhook	Bone	1276.4 ±382.9	465.7 ±111.5	132.5 ±23.8	0.7 ±0.4	0.0 ±0.0	2990	0.0114	88	0.987
2005/915	Bone	124868.6 ±12529.4	32242.3 ±2762.8	20090.5 ±6625.8	3169.4 ±396.2	749.4 ±79.6	170416	0.0080	125	0.978

Table S2: Sample characteristics

Collection numbers refer to the Cerpolex/Mammuthus Catalog.

Sample	¹⁴ C age	Discovery date	Preservation since then	Shaft sample	Root sample	Bone sample
Lyakhov	>50000 yBP	1902	Room temperature	31 mg (12 hair)	1 mg (8 roots)	110 mg
Jarkov (2003/811)	20380 ±140 yBP	1997	Frozen	54 mg (many hair)	2 mg (19 roots)	100 mg
FishHook (2001/384)	20620 ±70 yBP	1990	Frozen	-	-	110 mg
2005/915	27740 ±220 yBP	2005	Frozen	-	-	110 mg

Table S3: Total DNA yield of extracts as measured per PicoGreen Assay

Specimen	Sample type	Quantitation Reading ng.ml ⁻¹	Dilution factor	Extract total volume ml	Material extracted mg	Total [DNA] in sample ng.mg ⁻¹
Lyakhov	Root	2.38±0.01	50	0.1	1	11.91±0.05
	Shaft	2.87±0.21	50	0.1	31	0.46±0.03
	Bone	232.80±0.56	50	0.1	110	10.58±0.03
Jarkov	Root	2.66±0.00	50	0.1	2	13.30±0.00
	Shaft	4.19±0.40	50	0.1	54	0.39±0.04
	Bone	76.42±1.16	50	0.1	100	3.82±0.06
Fishhook	Bone	14.44±0.34	50	0.1	110	0.66±0.01
2005/915	Bone	43.31±0.32	50	0.1	110	1.97±0.01

Table S4: Primers listPrimer names indicate position on the *cytochrome b* sequence and direction (Forward/Reverse).

Primer	Sequence (5'-3')	T _m (°C)	Amplicon length (bp)
F111-131	AGGAGCATGCCTAATTACCCA	57	Variable
R171-194	GATGAAAATGCAGTTATTGTGTC	56	84
R241-261	TGCTCCGTTTGAGTGTAGTTG	55	151
R371-389	CCTATGAAGGCGGTGGCTA	57	279
R572-600	AAGGAAGGTTAGGTGTACTCCTGCTAGTG	60	490
R763-787	TAAGTGGATCAGCTGGTATGTAGTT	55	677

Table S5: qPCR raw output for standards and tested samples

Results organized per fragment size. Ct1 and Ct2 refer to the two replicates for each reaction (given per PCR reaction using 4 ul of extract). Standards (stds) were used to derive the standard curve slope, Y-intercept and R2. The efficiency of each assay was derived from its standard curve. Extractions blanks for shaft and root (S/R Blk) and bone (B Blk) as well as PCR Blanks were also run in duplicates.

		84 bp		151 bp		279 bp		490 bp		677 bp	
		Ct1	Ct2	Ct1	Ct2	Ct1	Ct2	Ct1	Ct2	Ct1	Ct2
stds	4000000	12,82	11,85	10,78	11,04	11,33	11,41	11,90	12,46	18,55	18,44
	400000	15,34	16,96	15,10	15,75	15,27	15,30	15,74	15,35	21,81	21,60
	40000	19,84	19,73	17,66	19,50	19,35	18,72	17,79	19,77	25,57	26,39
	4000	23,18	23,00	22,46	22,03	23,30	22,24	21,83	23,64	29,08	28,86
	400	25,90	25,54	24,03	26,57	25,88	26,52	25,12	27,60	33,66	33,53
	40	29,84	30,67	29,87	29,65	31,43	30,44	31,67	32,03	37,46	37,43
	4	-	-	33,47	32,28	33,81	34,82	-	-	39,82	38,89
std curve	slope	-3.4785		-3.6174		-3.8317		-3.8499		-3.6304	
	y-intercept	35,498		35,186		36,647		37,034		42,441	
	R2	0,9913		0,9898		0,9966		0,9786		0,9929	
	efficiency	0,9386		0,8899		0,8238		0,8186		0,8856	
samples	Lyakhov root	30,13	30,07	32,44	33,67	-	-	-	-	-	-
	Lyakhov shaft	31,71	31,31	33,15	-	-	-	-	-	-	-
	Lyakhov bone	22,20	22,83	24,80	25,16	-	-	-	-	-	-
	Jarkov root	25,40	24,12	25,93	27,10	32,81	33,99	-	-	-	-
	Jarkov shaft	21,93	21,81	23,93	23,70	29,95	29,69	35,50	-	-	-
	Jarkov bone	21,82	21,47	22,51	22,76	25,38	24,41	29,90	29,58	40,14	42,15
	Fishhook bone	26,29	25,64	26,58	27,12	29,68	30,11	39,96	38,52	-	-
	2005/915	18,91	19,12	20,07	20,27	21,97	21,18	25,08	24,79	33,10	33,32
blanks	S/R Blk	-	-	-	-	-	-	-	-	-	-
	B Blk	-	-	-	-	-	-	-	-	-	-
	PCR Blk	-	-	-	-	-	-	-	-	-	-

Table S6: Inhibition test

Percentage of recovery of copy numbers for serial dilutions of extracts organized by dilution factor (on top). Lyakhov and Jarkov hair extracts were compared based on different molecular weight ultrafiltration columns (see Materials and Methods): YM-10 or YM-30. The copy number for the lowest dilution which provided a positive product was arbitrarily assigned a 100% value for each sample. The percentages of recovery for the other dilutions were calculated from that value, according to the dilution factor. "No Ct" indicates that the qPCR yielded no product. Final dilutions used in the comparative analysis are in bold character for each extract assayed.

Microcon	Specimen	Sample type	straight	1:5	1:10	1:50	1:100	1:500	1:1000	1:5000
YM-10 extracts	Lyakhov	Root	No ct	-	No ct	No ct	No ct	No ct	No ct	-
		Shaft	No ct	-	No ct	No ct	No ct	No ct	No ct	-
	Jarkov	Root	No ct	-	No ct	12%	100%	100%	-	-
		Shaft	No ct	-	No ct	No ct	No ct	8%	71%	100%
YM-30 extracts	Lyakhov	Root	53%	-	100%	No Ct	-	-	-	-
		Shaft	No Ct	-	100%	100%	-	-	-	-
	Jarkov	Bone	32%	73%	89%	98%	100%	-	-	-
		Root	10%	-	100%	100%	-	-	-	-
		Shaft	No ct	-	No ct	15%	100%	100%	100%	-
	Fishhook 2005/915	Bone	No Ct	54%	90%	100%	100%	-	-	-
		Bone	0%	8%	45%	100%	100%	-	-	-
		Bone	31%	60%	93%	100%	100%	-	-	-

Table S7: Quantitation rescaling factor calculation

Dilution factors of the extracts are derived from Table S6 (a value of 1 corresponds to an extract assayed straight).

Specimen	Sample type	Extract total volume (ul)	PCR reaction volume (ul)	Dilution factor of the extract	Amount of material in the extract (mg)	Rescaling factor of copy numbers
Lyakhov	Root	100	4	1	1	25.000
	Shaft	100	4	10	31	8.064
	Bone	100	4	5	110	1.136
Jarkov	Root	100	4	10	2	125.000
	Shaft	100	4	100	54	46.296
	Bone	100	4	10	100	2.500
Fishhook 2005/915	Bone	100	4	10	110	2.273
	Bone	100	4	10	110	2.273

Table S8: Rate of damage for the bone sample 2005/915.

An amplicon of 788 bp that spans the 3' end of the cytochrome b and the Hypervariable Region I of the mitochondria was selected (4). Primers specific to elephantid sequences were designed. Two PCR products were generated from the 2005/915 bone extract. Eighteen clones were sequenced and compared to evaluate the number of miscoding lesions in each clone compared to the consensus sequence. Nucleotide composition of the 743 bp consensus sequence is as follows: A (244), G (95), C (181), T (223).

Clone	Apparent C > T damage	Apparent G > A damage	Other damage (apparent T > C)
1	1	0	0
2	1	0	0
3	0	0	0
4	1	0	0
5	1	0	0
6	0	0	1
7	1	0	0
8	0	0	1
9	1	0	0
10	0	0	0
11	3	0	0
12	2	0	0
13	0	0	1
14	3	0	0
15	2	0	0
16	0	2	0
17	3	2	0
18	1	0	1
frequency	0.403%	0.080%	0.048%

Supporting references

1. M. T. Gilbert *et al.*, *Science* **317**, 1927 (Sep 28, 2007).
2. H. N. Poinar *et al.*, *Science* **311**, 392 (2006).
3. B. E. Deagle, J. P. Eveson, S. N. Jarman, *Front Zool* **3**, 11 (2006).
4. R. Debruyne *et al.*, *Current Biology* **Accepted** (2008).