Letter to the Editor

Crosslinks Rather Than Strand Breaks Determine Access to Ancient DNA Sequences From Frozen Sediments

Anders J. Hansen,[*,1] **David L. Mitchell,**[†] **Carsten Wiuf,**[‡] **Lakshmi Paniker,**[†] **Tina B. Brand,**[*] **Jonas Binladen,**[*] **David A. Gilichinsky,**[§] **Regin Rønn**[**] **and Eske Willerslev**[*,1,2]

[*]Centre for Ancient Genetics, Niels Bohr Institute and Biological Institute, University of Copenhagen, DK-2100 Copenhagen Ø, Denmark, [†]Department of Carcinogenesis, University of Texas, M. D. Anderson Cancer Center, Science Park/Research Division, Smithville, Texas 7895, [‡]BiRC—Bioinformatics Research Center, University of Aarhus, 8000 Aarhus C, Denmark, [§]Soil Cryology Laboratory, Institute for PhysicoChemical and Biological Problems in Soil Science, Russian Academy of Sciences, 142290, Pushchino, Moscow Region, Russia and [**]Department of Terrestrial Ecology, Biological Institute, University of Copenhagen, DK-1353 Copenhagen K, Denmark

ABSTRACT

Diagenesis was studied in DNA obtained from Siberian permafrost (permanently frozen soil) ranging from 10,000 to 400,000 years in age. Despite optimal preservation conditions, we found the sedimentary DNA to be severely modified by interstrand crosslinks; single- and double-stranded breaks; and freely exposed sugar, phosphate, and hydroxyl groups. Intriguingly, interstrand crosslinks were found to accumulate ~100 times faster than single-stranded breaks, suggesting that crosslinking rather than depurination is the primary limiting factor for ancient DNA amplification under frozen conditions. The results question the reliability of the commonly used models relying on depurination kinetics for predicting the long-term survival of DNA under permafrost conditions and suggest that new strategies for repair of ancient DNA must be considered if the yield of amplifiable DNA from permafrost sediments is to be significantly increased. Using the obtained rate constant for interstrand crosslinks the maximal survival time of amplifiable 120-bp fragments of bacterial 16S ribosomal DNA was estimated to be ~400,000 years. Additionally, a clear relationship was found between DNA damage and sample age, contradicting previously raised concerns about the possible leaching of free DNA molecules between permafrost layers.

THE successful retrieval of ancient DNA (aDNA) sequences from fossil remains is limited by the postmortem instability of nucleic acids. It is generally assumed that DNA fragmentation caused by depurination (followed by cleavage of the phosphodiester backbone by β-elimination) is the fastest of the spontaneous chemical reactions limiting the half-life of amplifiable aDNA (*e.g.*, PÄÄBO and WILSON 1991; POINAR *et al.* 1996, 2002; MAROTA *et al.* 2002). However, the notion behind this assumption remains untested, and some recent studies suggest that crosslinking preventing amplification due to reduced denaturation might be a significant contributor to DNA diagenesis in some environments (POINAR *et al.* 1996; WILLERSLEV *et al.* 2004a; WILLERSLEV and COOPER 2005). Here we address the role of strand breaks *vs.* crosslinks in limiting the amplification of aDNA molecules under permafrost conditions. This was done by comparing the relative rates of single-stranded breaks (SSBs) and interstrand crosslinks (ICLs) in newly obtained DNA and previously described DNA (MITCHELL *et al.* 2005) extracted directly from five Siberian permafrost cores ranging from 10,400 to 400,000–600,000 years in age (for experimental details see supplemental materials at http://www.genetics.org/supplemental/).

Depurination and crosslinking: The presence of depurination-generated SSBs was investigated by treating parts of the extracted permafrost DNA with alkali or shrimp alkaline phosphatase (SAP). These treatments remove sugar residues and phosphate groups, respectively, leaving 5′-termini OH groups that are available for a radioactive phosphate group (^{32}P), using the T4 polynucleotide kinase (PNK) enzyme (SAMBROOK and RUSSEL 2001). Additionally, radioactive nucleotides were incorporated into the recessed or intrinsic 3′-OH ends of the permafrost DNA, using the Klenow fragment of *Escherichia coli* DNA pol I (TELFORD *et al.* 1979). These treatments revealed the presence of a substantial number of

---

[1]These authors contributed equally to this work.

[2]*Corresponding author:* University of Copenhagen, Juliane Maries vej 30, DK-2100 Copenhagen Ø, Denmark.   E-mail: ewillerslev@bi.ku.dk

TABLE 1

Lesion frequencies (f_{lesion}), rate constants (k), and half-lives ($T_{1/2}$)

Lesion type	Time (YBP)	f_{lesion}	k (sec^{-1})	$T_{1/2}$ (yr)
DSB[a]	10,400 → 300,000–400,000	0.00013	3.7×10^{-17}	8×10^8
	10,400 → 400,000–600,000	0.00037		
SSB[b]	10,400	0.00053	1.4×10^{-15}	1.7×10^7
	19,000	0.0009		
ICL[c]	10,400	0.44		
	19,000	0.49	1.0×10^{-13}	2.2×10^5
	300,000–400,000	0.85		
	400,000–600,000	0.87		

YBP, years before present.

[a] Lesion frequencies of double-stranded breaks (DSBs) were obtained by comparing number average molecular length (NAML) of double-stranded DNA from two sample pairs of different ages (10,400 *vs.* 300,000–400,000 YBP and 10,400 *vs.* 400,000–600,000 YBP).

[b] Lesion frequencies of single-stranded breaks (SSBs) were obtained by comparing NAML of double-stranded and single-stranded DNA from the same samples (10,400 and 19,000 YBP).

[c] Lesion frequencies of interstrand crosslinks (ICLs) were calculated on the basis of data from a PicoGreen assay on samples dated to be 10,400, 19,000, 300,000–400,000, and 400,000–600,000 years old. From the lesion frequencies rate constants were calculated followed by calculations of half-lives. See text and supplemental material (http://www.genetics.org/supplemental/) for details.

5′-OH termini, 5′-termini free sugar radicals, and 5′-PO₄ groups as well as 3′-OH termini (Wilcoxon matched pairs, $P < 0.05$), corresponding to the end products expected by SSBs generated after depurination by β-elimination (LINDAHL and NYBERG 1972). Additionally, the treatments showed higher yields of PO₄ groups relative to free sugar residues, agreeing with the pattern previously observed in depurination studies of fully hydrated DNA from phage PM2 (LINDAHL and ANDERSSON 1972). We therefore concluded that SSBs generated by depurination are present in the permafrost-preserved DNA.

The presence of ICLs was investigated by the refractivity of the DNA extracts to heat denaturation, using a very sensitive microplate fluorescence assay and Pico-Green (Molecular Probes, Eugene, OR) as described in WILLERSLEV *et al.* (2004a). The fluorescent DNA probe has 10-fold greater affinity for double-stranded than for single-stranded DNA, and thus a loss of duplex DNA induced by denaturation causes a significant reduction in fluorescence signal. By comparing the extent of signal reduction in the permafrost DNA with that of an undamaged control (*E. coli*) measured at comparable concentrations we found that all DNA extracts are significantly affected by ICLs (Table 1).

Rate constants: Lesion frequencies were obtained as described previously (LINDAHL and NYBERG 1972) and were used to calculate the rate constants for SSBs and ICLs by assuming that the DNA degradation follows first-order kinetics (Table 1). Thus, the amount of undamaged bases (B_t) at time t is given by the equation

$$B_t = B_0 e^{-kt},$$

where B_0 is the initial concentration of bases and k is the rate constant. The amount of damaged bases $B_{\text{d}} = B_0 - B_0 e^{-kt}$ and the lesion frequency ($f_{\text{lesion}} = B_{\text{d}}/B_0$) is given by

$$f_{\text{lesion}} = (B_0 - B_0 e^{-kt})/B_0 = 1 - e^{-kt}.$$

Since $e^{-kt} = 1 - f_{\text{lesion}}$, it follows that k can be estimated as the slope of the line $-\ln(1 - f_{\text{lesion}})$ plotted against time. This gave a rate constant for SSBs of 1.4×10^{-15} sec^{-1} (resulting in a half-life of 1.7×10^7 years; Table 1), which is somewhat similar to the depurination rate constant determined by LINDAHL and NYBERG (1972) of 2.5–3.1×10^{-15} sec^{-1} (extrapolated to $-12°$, the temperature of the permafrost). The rate constant for ICLs was calculated to be 1.0×10^{-13} sec^{-1}, which is ∼40 times faster than the preliminary estimates reported by MITCHELL *et al.* (2005), and results in a half-life of 2.2×10^5 years. This suggests that ICLs are accumulating ∼100 times faster than SSBs under frozen conditions (Table 1). As crosslinks hinder DNA amplification by preventing template denaturation, the result contradicts the general assumption that fragmentation through depurination is the most important type of damage limiting the amount of amplifiable DNA in fossil remains, including those in permafrost. This finding is in agreement with previous claims of measurable putative DNA sequences being present in permafrost even though amplification products cannot be obtained (WILLERSLEV *et al.* 2004a).

Predicting DNA survival: Some of the most commonly used models for predicting the long-term survival of DNA in fossil remains are those based on amino acid
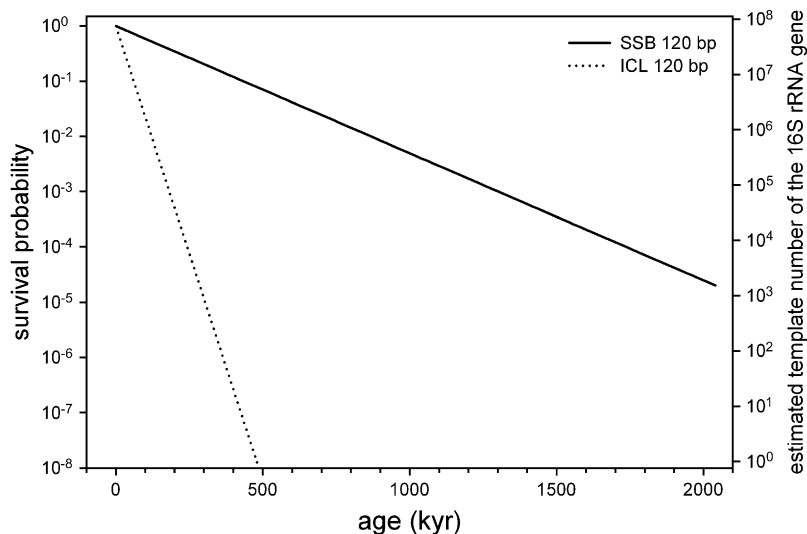
Figure 1.—The survival probability of 120-bp DNA fragments in Siberian permafrost estimated from rate constants for single-stranded breaks and interstrand crosslinks. The secondary y-axis shows the corresponding number of remaining templates of bacterial 16S rRNA genes that would be present in one PCR reaction (as performed in Willerslev *et al.* 2004a) assuming an initial bacterial density of 10^8 g^{-1} of permafrost soil when entering the permanently frozen state. The probability of fragments avoiding interstrand crosslinks was estimated from the rate constant estimated by the PicoGreen assay by assuming that each crosslink affects the affinity for Pico-Green in a 10-bp part of the double-stranded DNA. For further details see text.

racemization (Poinar *et al.* 1996) and thermal age (Smith *et al.* 2001). Implicit in these models is a belief that depurination is the only significant type of damage limiting the half-life of amplifiable aDNA. Recent studies of DNA from fossil coprolites that are preserved under warm and arid conditions have questioned the general validity of this assumption by showing that intermolecular crosslinks, rather than fragmentation, determine the access to template molecules for PCR (Poinar *et al.* 1998). Our results suggest that also under frozen conditions the key factor in determining the half-life of amplifiable DNA is crosslinks rather than SSBs. Thus, the obtained results seriously question the reliability of models purely relying on depurination kinetics to predict the long-term survival of permafrost-preserved DNA. Importantly, Arctic permafrost is a fairly dry environment with only 3–8% of the total water being in an unfrozen state (Willerslev *et al.* 2004b) and limited hydrolytic damage could possibly account for the fairly low levels of observed SSBs relative to crosslinks. Additionally, the close to neutral pH in the Siberian permafrost (Willerslev *et al.* 2004a) favors long-term DNA preservation. Thus, similar studies are needed on various types of materials to determine to what extent these results can be generalized. In our opinion, it is likely that the relative rates of various damage types are highly variable across environments, depending on their exact composition of free water, pH, salt, oxidative free radicals, etc. Therefore, it is most likely impossible to use general models, such as those of amino acid racemization (Poinar *et al.* 1996) and thermal age (Smith *et al.* 2001), to predict the long-term survival of DNA across all environments. Instead, damage rates for each unique environment need to be investigated before more realistic predictions on DNA survival, such as those shown below, can be made.

Bacterial DNA survival: The estimated rate constants for the degradation of DNA in permafrost samples permit us to roughly calculate, for a given starting number of template molecules, the expected number remaining after a given time. Assuming that the DNA degradation follows first-order kinetics, the amount of undamaged bases at time t is $B_t = B_0 e^{-kt}$, and the proportion of undamaged bases is $B_t/B_0 = e^{-kt}$. Under the assumption that all bases have the same chance of being damaged, this corresponds to the probability that any given base is still undamaged at time t. The chance that two neighboring bases are both undamaged at time t is then $(e^{-kt})^2$ assuming independence, and the chance that all bases in a fragment of length n are undamaged is $(e^{-kt})^n$. From this we first estimated the probability of survival of 120-bp fragments, assuming that the only process that affects DNA survival is single-stranded breaks (Figure 1). The survival probability decreases rapidly and after 2 million years the chance of a 120-bp fragment surviving would be 2.5×10^{-5}. Although this probability is low, in light of the high initial number of templates that would have been present when the soil entered the frozen state, this still would result in the survival of a considerable number of template molecules. For example, direct counts of bacteria in surface soils usually reveal $\sim 10^9$ bacterial cells per gram of soil, and a specific study from Siberian permafrost found 10^8 bacteria in the active layer directly above the permanently frozen soil (Kobabe *et al.* 2004). In a previous study, we attempted to amplify bacterial DNA fragments from the same permafrost samples as used in this study, using DNA extracted from 2 g of soil (Willerslev *et al.* 2004a). If we assume that each of the bacterial cells had, on average, 3.8 copies of the 16S rRNA gene (Fogel *et al.* 1999), then one PCR reaction from the 1.5- to 2-million-year-old sample would contain between 2.0×10^3 and 2.6×10^4 templates [3.8 copies cell^{-1} $\times$ 2 g $\times$ 10^8 cells g^{-1} (5 μl/50 μl) (2.5×10^{-5}) and 3.8 copies cell^{-1} $\times$ 2 g $\times$ 10^8 cells g^{-1} (5 μl/50 μl) (3.5×10^{-4})] (Figure 1). This number should be sufficient to reproducibly allow

for amplification of template DNA (Handt *et al.* 1996). However, in the previous study we were unable to amplify bacterial 16S rDNA fragments of 120 bp from DNA extracts older than 400,000–600,000 years (Willerslev *et al.* 2004a), suggesting that processes other than SSBs are involved.

It is more complicated to use the rate constants for interstrand crosslinks to make similar calculations. The PicoGreen assay does not provide exact information on how single base pairs behave. The affinity for PicoGreen of double-stranded DNA is probably increased some distance from a crosslink. Even if we assume that each crosslink results in increased affinity up to 5 bp during denaturation, crosslinks would reduce amplifiability much more rapidly than single-stranded breaks. Accordingly, amplifiable 120-bp bacterial 16S rDNA should not be present after ~400,000 years (Figure 1). This estimate is in agreement with our previous results showing independent and reproducible amplification of 120-bp 16S bacterial rDNA from the permafrost samples dated to be up to 400,000–600,000 years old (the same samples as used in this study), but not from samples dated to be 1.5–2 million years old (Willerslev *et al.* 2004a). These simple calculations additionally support the view that ICLs rather than SSBs are the most important damage type prohibiting amplification of DNA extracted from the permafrost samples.

Damage/age relationships and DNA leaching: Although theory predicts an increase in DNA damage with time (Hebsgaard *et al.* 2005), previous studies of diagenesis in fossil remains have not been able to confirm this (Pääbo 1989; Höss *et al.* 1996; Gilbert *et al.* 2003). A likely explanation is that preservation conditions rather than the age of the specimens determine the degree of DNA degradation (Willerslev *et al.* 2004b). Interestingly, this is the first time that DNA damage can be clearly linked to sample age. This is true for strand breaks and ICLs (Table 1). Additionally, statistical analysis showed a significant increase (ANOVA, $P = 0.025$) in incorporation of radioactive phosphates (*i.e.*, the PNK end-labeling assays after alkaline or SAP treatments) with sample age, as did the Klenow fill-in of radioactive nucleotides (Wilcoxon matched pairs, $P < 0.05$). These observations are in agreement with the geological interpretation that Siberian permafrost preservation conditions must have remained constant over time, making permafrost unique for comparative studies of DNA damage rates. Furthermore, these observations provide further support of the authenticity of the results, as contaminated samples would not be expected to produce such an age-specific relationship. Such a relationship between damage and age is unlikely to be found in most other environments due to variable preservation conditions. Importantly, the observed damage/age relationship reveals that leaching of DNA between sediment layers must be minor at most. The sections under study do not constitute a continuous layering of sediments

due to, among other things, tectonic uplift and erosion. Therefore, leaching of DNA between layers would disturb any clear pattern between the sediment age and damage of the DNA. The results therefore contradict previously raised concerns about the leaching of free DNA molecules between frozen layers that would severely complicate any paleoenvironmental reconstructions based on permafrost-preserved sedimentary DNA (Pääbo *et al.* 2004).

Ancient DNA repair: Limited success of aDNA repair has been achieved using DNA polymerase I (*E. coli*), and T4 DNA ligase (Pääbo 1989; Pusch *et al.* 1998). Our results confirm that this approach will probably repair some of the gaps and nicks in the ancient sediment DNA. However, as a major part of the 5′ termini do not have PO_4 groups, and as this method requires the presence of 3′-OH and 5′-PO_4 termini, the amount of repairable ancient templates must be limited. Furthermore, because SSBs in the aDNA most likely result from depurination, which consequently results in 3′ termini with aldehyde groups, gap filling by DNA polymerase will be inhibited. To repair the majority of SSBs in aDNA we propose a pretreatment with PNK and AP endonuclease I, which will assure that most 5′ and 3′ ends have PO_4 and OH groups, respectively. Subsequent to this initial enzymatic treatment, gap filling by a polymerase with a low strand displacement capability and 3′ → 5′ exonuclease activity (*e.g.*, *E. coli* Pol II), which can remove carbon radicals from the 5′ end, and gap filling by T4 DNA ligase should greatly increase the yield of amplifiable aDNA fragments. A note of caution, however, is that applying enzymatic assays like the ones described to ancient DNA is very difficult. Many of the repair techniques need to be optimized to very low amounts of DNA. Issues like contamination with modern templates during the experimental procedure need to be dealt with efficiently if the end product of the repaired amplified sequence is to be reliable and credible. Contaminating DNA molecules can enter the procedure at different levels, so appropriate controls and criteria need to be developed (Willerslev and Cooper 2005).

LITERATURE CITED

Fogel, G. B., C. R. Collins, J. Li and C. F. Brunk, 1999 Prokaryotic genome size and SSU rDNA copy number: estimation of microbial relative abundance from a mixed population. Microb. Ecol. **38:** 93–113.

Gilbert, M. T. P., A. J. Hansen, E. Willerslev, I. Barnes, L. Rudbeck *et al.*, 2003 Characterisation of genetic miscoding

lesions caused by post-mortem damage. Am. J. Hum. Genet. **72:** 48–61.

HANDT, O., M. KRINGS, R. H. WARD and S. PÄÄBO, 1996 The retrieval of ancient human DNA sequences. Am. J. Hum. Genet. **59:** 376–386.

HEBSGAARD, M. B., M. PHILLIPS and E. WILLERSLEV, 2005 Geologically ancient DNA: Fact or artefact? Trends Microbiol. **13:** 212–220.

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

KOBABE, S., D. WAGNER and E. PFEIFFER, 2004 Characterisation of microbial community composition of a Siberian tundra soil by fluorescence in situ hybridisation. FEMS Microbiol. Ecol. **50:** 13–23.

LINDAHL, T., and A. ANDERSSON, 1972 Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid. Biochemistry **11:** 3618–3623.

LINDAHL, T., and B. NYBERG, 1972 Rate of depurination of native deoxyribonucleic acid. Biochemistry **11:** 3610–3618.

MAROTA, I., C. BASILE, M. UBALDI and F. ROLLO, 2002 DNA decay rate in papyri and human remains from Egyptian archaeological sites. Am. J. Phys. Anthropol. **117:** 310–318.

MITCHELL, D., E. WILLERSLEV and A. J. HANSEN, 2005 Damage and repair of ancient DNA. Mutat. Res. **571:** 265–276.

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

PÄÄBO, S., and A. C. WILSON, 1991 Miocene DNA sequences—A dream come true? Curr. Biol. **1:** 45–46.

PÄÄBO, S., H. POINAR, D. SERRE, V. JAENICKE-DESPRES, J. HEBLER et al., 2004 Genetic analyses from ancient DNA. Annu. Rev. Genet. **38:** 645–679.

POINAR, H., 2002 The genetic secrets some fossils hold. Acc. Chem. Res. **35:** 676–684.

POINAR, H. N., M. HÖSS, J. L. BADA and S. PÄÄBO, 1996 Amino acid racemization and the preservation of ancient DNA. Science **272:** 864–866.

POINAR, H. N., M. HOFREITER, G. S. SPAULDING, P. S. MARTIN, A. B. STANKIEWICZ et al., 1998 Molecular coproscopy: dung and diet of the extinct ground sloth Nothrotheriops shastensis. Science **281:** 402–406.

PUSCH, C., I. GIDDINGS and M. SCHOLZ, 1998 Repair of degraded duplex DNA from prehistoric samples using Escherichia coli DNA polymerase I and T4 DNA ligase. Nucleic Acids Res. **26:** 857–859.

SAMBROOK, J., and D. W. RUSSEL, 2001 Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

SMITH, C. I., A. T. CHAMBERLAIN, M. S. RILEY, A. COOPER, C. B. STRINGER et al., 2001 Neanderthal DNA: Not just old but old and cold? Nature **10:** 771–772.

TELFORD, J. L., A. KRESSMANN, R. A. KOSKI, R. GROSSCHEDL, F. MULLER et al., 1979 Delimitation of a promoter for RNA polymerase III by means of a functional test. Proc. Natl. Acad. Sci. USA **76:** 2590–2594.

WILLERSLEV, E., and A. COOPER, 2005 Ancient DNA. Proc. Biol. Sci. **272:** 3–16.

WILLERSLEV, E., A. J. HANSEN, T. B. BRAND, R. RØNN, I. BARNES et al., 2004a Long-term persistence of bacterial DNA. Curr. Biol. **14:** R9–R10.

WILLERSLEV, E., A. J. HANSEN and H. N. POINAR, 2004b Isolation of nucleic acids and cultures from ice and permafrost. Trends Ecol. Evol. **19:** 141–147.

Communicating editor: N. TAKAHATA