Available online at www.sciencedirect.com
SCIENCE DIRECT®
Mutation Research 571 (2005) 265–276
Fundamental and Molecular Mechanisms of Mutagenesis
www.elsevier.com/locate/molmut
Community address: www.elsevier.com/locate/mutres

Review

Damage and repair of ancient DNA

David Mitchell [a,*], Eske Willerslev [b], Anders Hansen [c]

[a] *The University of Texas MD Anderson Cancer Center, Department of Carcinogenesis, P.O. Box 389, Park Road 1C, Smithville, TX 78957, USA*
[b] *Henry Wellcome Ancient Biomolecules Centre, Department of Zoology, University of Oxford, Oxford, UK*
[c] *Department of Evolutionary Biology, Zoological Institute, University of Copenhagen, Copenhagen, Denmark*

Received 3 June 2004; received in revised form 25 June 2004; accepted 25 June 2004

Abstract

Under certain conditions small amounts of DNA can survive for long periods of time and can be used as polymerase chain reaction (PCR) substrates for the study of phylogenetic relationships and population genetics of extinct plants and animals, including hominids. Because of extensive DNA degradation, these studies are limited to species that lived within the past 10^4–10^5 years (Late Pleistocene), although DNA sequences from 10^6 years have been reported. Ancient DNA (aDNA) has been used to study phylogenetic relationships of protists, fungi, algae, plants, and higher eukaryotes such as extinct horses, cave bears, the marsupial wolf, the moa, and Neanderthal. In the past few years, this technology has been extended to the study of infectious disease in ancient Egyptian and South American mummies, the dietary habits of ancient animals, and agricultural practices and population dynamics of early native Americans. Hence, ancient DNA contains information pertinent to numerous fields of study including evolution, population genetics, ecology, climatology, medicine, archeology, and behavior. The major obstacles to the study of aDNA are its extremely low yield, contamination with modern DNA, and extensive degradation. In the course of this review, we will discuss the current aDNA literature describing the importance of aDNA studies as they relate to important biological questions and the difficulties associated with extracting useful information from highly degraded and damaged substrates derived from limited sources. In addition, we will present some of our own preliminary and published data on mechanisms of DNA degradation and some speculative thoughts on strategies for repair and restoration of aDNA.

Keywords: Ancient DNA; DNA fragmentation; DNA repair; Interstrand cross-links; Oxidative damage

Abbreviations: aDNA, ancient DNA; ALS, alkali-labile site; AP, apurininc/apyrimidinic; B.P., before present; Dsb, double-strand break; GC/MS, gas chromatography/mass spectrometry; ICL, interstrand cross-link; mtDNA, mitochondrial DNA; NAML, number average molecular length; PCR, polymerase chain reaction; PNK, polynucleotide kinase; SAP, shrimp alkaline phosphatase; Ssb, single-strand break

* Corresponding author. Tel.: +1 512 237 9474.
E-mail address: dmitch@starband.net (D. Mitchell).

0027-5107/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.mrfmmm.2004.06.060

Contents

The past 15 years has seen a burgeoning of studies utilizing ancient DNA (aDNA) technologies to study a variety of biological problems pertinent to numerous fields of study including evolution, population genetics, ecology, climatology, medicine, molecular archeology and anthropology, cultural evolution, behavior, and exobiology to name a few [1–4]. The impact of this technology has been significant and its future holds considerable promise. This review is designed to give a brief summary, with specific examples, of aDNA insights into paleobiology and human prehistory, outline some of the problems associated with aDNA analysis with specific focus on DNA degradation and damage accumulation, and discuss the prospect of repairing aDNA and increasing the amount of sequence information available for polymerase chain reaction (PCR) analysis. Some of our recent data on aDNA damage will be summarized as well.

1. Molecular paleobiology

A major application of aDNA technology has been in the fields of evolution and population genetics. DNA extracted from a North Greenland ice core dated at 2000 and 4000 years B.P. (before the present) revealed a high diversity of fungi, plants, algae, and protists derived from the local environment as well as distant sources [5]. Similarly, aDNA extracted from five permafrost cores from Siberia, ranging from 10,000 to 400,000 years B.P., revealed a broad range of taxonomic diversity in the paleoenvirnoment [6]. Sequences were identified for 19 different plant taxa as well as several mammalian species including mammoth, bison, and horse and showed significant changes in the diversity and composition of plants and animals across the late Pleistocene ice-free zone that stretched across Northeast Siberia and the Bering land bridge to Western Canada. Phylogenetic relationships have also been explored in populations of extinct horses [7], cave bears [8], ground sloths [9], the marsupial wolf [10], saber-toothed cats [11], the wooly mammoth [12], the moa [13] and most recently, Adelie penguins [14], to name a few. Recent technology has allowed DNA extraction from coprolites, fossilized fecal material, which contain sequences from both the host animal and its diet. From this material investigations on the dietary habits of ancient animals, such as the ground sloth have provided evidence for ecological and climate change in the distant past [15].

Ancient DNA analyses have also been applied to problems of human history and prehistory and promise to add invaluable and exciting insights into these fields in the near future. Hybridization of genomic DNAs isolated from modern humans with that from parietal and clavicle fossils of *Homo neanderthal* from Germany and Croatia, respectively, showed considerably more divergence than similar analyses using DNA from early *Homo sapiens* remains from Vogelherd cave in Germany. These results support the idea that Neanderthal was not in the direct lineage of modern man [16]. A similar study using mitochondria DNA from a 29,000-year-old Neanderthal specimen recovered from Mezmaiskaya Cave in the Northern

Caucasus further support this idea as well as the idea that modern human evolution was not multiregional [17]. These and other molecular studies on early man are controversial, plagued by the extremely small yields of aDNA, the related problems of modern human DNA contamination during the extraction process, as well as the extensive degradation associated with long-term storage in suboptimal conditions. These issues will be discussed below.

Ancient DNA analyses have also been important in the study of human specimens from the ancient historical past. A recent study took advantage of Victorian anthropological collections from the Natural History and British museums in London to examine the origins of a distinct and extinct population of humans who lived in the Andaman Archipelago in the Bay of Bengal [18]. In another study, mitochondrial DNA (mtDNA) sequences from 27 prehistoric Anasazi specimens from the Southwest United States suggest that these populations were much more stable than previously thought after European colonization and may be associated with modern Pueblo groups [19].

In the past few years, a variety of bacterial, protozoal, and viral infections have been detected in ancient mummified and skeletal tissues using PCR amplification and sequencing [20]. One particular paleopathology study identified human tuberculosis in bone tissue samples from three different populations in ancient Egypt ranging from predynastic (as early as 3500 B.C.) to the late period (as late as 500 B.C.) [21]. Sequence data correlated the presence of mycobacterial aDNA with pathologies associated with tuberculosis but also showed widespread occurrence of this pathogen in non-infected individuals. *Mycobacterium tuberculosis* has also been detected in tissues from mummies found in the Andes Mountains dating from 140 to 1200 A.D. [22]. Recent results using DNA extracted from a broad temporal range of dessicated human mummies from coastal and low valley sites in Northern Chile and Southern Peru indicate that the parasite *Trypanosoma cruzi* associated with Chagas' disease was prevalent at the time the earliest humans populated this area of South America [23]. DNA extracted from coprolites have also been used to offer a unique glimpse at the dietary and, perhaps, agricultural behavior of native Americans living 2000 years ago in Nevada [24].

Because of extensive DNA degradation, all of these studies are limited to species that lived within the past 10^5 years (Late Pleistocene). Although DNA sequences from 10^6 years have been reported [3,25] the results are not reproducible and the data are suspect. The major obstacles to the study of aDNA are its extremely low yield, contamination with modern DNA, and extensive degradation. With few exceptions (e.g., see [26]), most of the studies to date have focused on mtDNA due to the presence of multiple copies per cell and an increased likelihood of PCR amplification. However, because mtDNA data represent a single-genetic locus, the results may not reflect population trends in the genome overall. Although preferable, PCR analysis of nuclear DNA is more difficult since very little survives in ancient tissues and what does survive is badly damaged. Compounding these problems is the presence of modern DNA contaminants that need only occur in extremely small amounts to effectively compete for PCR amplification. This is particularly germane to the study of ancient human DNA where small amounts of contemporary DNA may be indistinguishable from the DNA of interest and are the most likely surface contaminant.

2. The chemistry of DNA degradation

DNA is a relatively unstable biological molecule and the damage that can accumulate over extended periods of time in "preserved" specimens is enormous and can include a wide diversity of lesions. Early work by Paabo [27] examined the integrity of DNA extracted from the dry remains of soft tissues ranging from 4 to 13,000 years old, including the extinct marsupial wolf and giant ground sloth. The purified DNA was low molecular weight and contained extensive damage including modified pyrimidines and sugar residues, abasic sites, interstrand cross-links (ICLs), and deamination products. Although much of the damage likely results from spontaneous degradation, a significant portion may arise from environmental exposures such as low level radiation or genotoxic chemicals (DNA damaging agents). Indeed, damage can also be introduced during the DNA extraction and purification processes.

Cleavage of the phosphate sugar backbone, a single-strand break (Ssb), is the predominant type of damage in aDNA resulting from direct hydrolysis or destabi-

lization at sites of base loss. The instability and decay of the primary structure of DNA was investigated 30 years ago using ^{14}C-labeled purine and pyrimidine residues to measure the rates of release of free bases as a function of temperature, pH, ionic strength, and nucleic acid secondary structure [28–31]. The results showed that purine bases were lost ~20-fold faster than pyrimidine bases. From the slope of the Arrhenius plot generated (rate of depurination versus temperature), Lindahl and Nyberg [28] calculated the activation energy of depurination at 31.2 kcal/mole, which translates to the loss of about 0.5 purines/*Escherichia coli* genome/generation (i.e., 40 min at 37 °C). Because a linear correlation between base loss and temperature was determined over a wide range of temperatures it "appears implausible for thermodynamic reasons that a 'threshold' temperature exists below which depurination of DNA would not occur." Using the activation energy, we calculated a 1000-fold decrease in the rate of depurination between 37 and 0 °C and a 350,000-fold decrease as the $T°$ drops from 0 to −50 °C. At sites of base loss, the aldehyde form of the deoxyribose sugar is vulnerable to cleavage by β-elimination leaving a degraded sugar residue and DNA fragments with terminal 3′ and 5′ phosphate groups. The average lifetime of the phosphodiester bond at an apurinic or apyrimidinic (AP) site is 10 h at 37 °C [32].

In addition to the intrinsic destabilization of DNA structure by labile glycosylic bonds, DNA bases are susceptible to hydrolytic deamination. Deamination rates of ^{14}C-labeled cytosine and 5-methylcytosine were determined experimentally by following their conversion to uracil and thymine residues as a function of temperature, pH, and DNA secondary structure [33–35]. In contrast to depurination, the double helix affords very good protection against hydrolytic cytosine deamination. Hence, although the rates of deamination and depurination are similar in single-stranded DNA, the rate of cytosine deamination in double-stranded DNA is only 0.5–0.7% the rate of depurination. Extrapolation of an Arrhenius plot to 0 °C gives a rate of cytosine deamination of 0.1–0.2 events in 10^6 bases DNA in 1 year; the less common base, 5-methylcytosine is deaminated three- to four-fold faster than cytosine. In addition to its temperature dependence, deamination is also pH dependent, occurring much more frequently under acid conditions.

3. Ancient DNA damage

As mentioned, the thermal environment in which aDNA samples are found is extremely important with respect to DNA degradation and damage [36]. Viable bacteria have been extracted from Arctic and Antarctic ice sheets as well as in mountain valley glaciers. Arctic ice collected from Ellesmere Island [37] identified many coliform bacteria some of which were dated at >2000-year-old and our data show a broad phylogenetic distribution of bacteria in Greenland ice (in preparation). Abyzov et al. from the Soviet Union conducted much of the work on the microbiology of ancient Antarctic ice [38–42]. Five strains of *Pseudomonas* bacteria were found in Central Antarctica near Vostok Station at depths of 79–81 and 91–92 m, corresponding to 2100 and 2500 years, respectively [43], and a strain of actinomycete was isolated from a layer of Vostok ice dated at 47,000 years [41]. Ice samples extracted from core depths of 1500 (110,000 years) and 2750 m (240,000 years) from Vostok hole Nr. 5G showed diverse microbial life ranging from 0.8×10^{-3} to 11×10^{-3} cells/mL. Although bacteria dominate the viable microorganisms extracted from core ice dated at 10–13,000 years [38] fewer are found in core samples from 40,000 years [39] and at depths dating to 150,000 years the numbers are extremely low [40]. Abyzov et al. compared microscopic cell counts with growth on nutrient media and consistently found that as the depth of the core increased, there was a gradual reduction in the number of viable bacteria [38–40]. From these studies, it appears that bacteria accumulate from a constant and fairly rapid drift of dust and debris onto the surface of the continental ice sheet and are transferred down into the interior of the ice sheet by continued deposition and accretion. In addition to bacteria, other organisms have been detected in glacial ice, including various marine microorganisms (e.g., diatoms) as well as fungal and algal spores [44]. Although the rate of DNA degradation is very slow at temperatures <0 °C, over extended periods of time significant amounts of damage will accumulate. The inverse correlation between recovery of viable bacteria and depth (i.e., age) in glacial ice may, in part, reflect the accumulation of endogenous DNA damage by glacial bacteria.

In contrast to glacial ice, organisms found in permafrost are for the most part endemic and include a high diversity of microbial, plant, and animal species

[44,45]. Indeed, permafrost is an important source of material for the evolutionary and phylogenetic studies discussed above. Permafrost is defined on the basis of temperature, as soil or rock that remains below 0 °C throughout the year, and forms when the ground cools sufficiently in winter to produce a frozen layer that persists throughout the following summer. Permafrost underlies an estimated 20–25% of the world's land surface; it occurs in more than 50% of Russia and Canada, 82% of Alaska, 20% of China, and much of Antarctica. Permafrost thickness ranges from 1600 m in Northern Siberia to 650 m in Northern Alaska. The active surface layer of permafrost is a thin slice of tundra vegetation that thaws every summer and freezes in winter.

Because of the low temperatures DNA from ice and permafrost, albeit degraded, survives for long periods of time, at least 50,000 years in glacial ice floes and considerably longer in permafrost. Although spontaneous DNA damage rates in glacial ice and permafrost may be similar, it is probable that very different environmental (chemical) conditions determine different DNA damage spectra. Once deposited, glacial DNA should not accumulate any significant environmental damage but should reflect only entropic, spontaneous effects on DNA structure. Although it is very doubtful that bacteria are exposed to genotoxic chemicals in glacial ice, it is possible that they are already damaged at the time of deposition on the surface of the glacier. The species diversity seen in various Arctic and Antarctic cores suggests that they may have traveled considerable distances before deposition and it is possible that aeolian exposure to UV radiation at high altitude may introduce DNA damage, some of which is indistinguishable from damage that forms spontaneously. The much richer organic composition of permafrost soils increases the likelihood that the DNA will accumulate significant levels of environmental damage. Some of the environmental damage will be similar to spontaneous damage, particularly when damage formation involves free-radical-mediated pathways (e.g., base loss and strand scission). However, other types of environmental damage may also be induced, including oxidations, alkylations, and cross-links resulting from exposure to direct-acting alkylating and cross-linking compounds (e.g., methyl chloride and nitrogen mustards). Low-level ionizing radiation may also make a noticeable impact on DNA degradation.

In a seminal paper published in 1989, Paabo [27] examined DNA damage and degradation in the dry remains of soft tissues ranging from 4 to 13,000 B.P. The DNA yields from these samples were relatively robust ranging from 1 to 200 μg/g of tissue (which ranged from 0.1 to 0.5 g). The DNA in all of the samples showed similar levels of degradation on agarose gels with molecular sizes ranging from 40 to 500 base pairs. HPLC analysis showed a severe loss of pyrimidine residues in all but the contemporary (4 years) sample and a corresponding increase in peaks possibly associated with ring-fragmented and/or ring-saturated pyrimidine derivatives. To further examine base damage in aDNA, Paabo measured the loss of TCA precipitable DNA (<20 bp) after treatment with alkali or digestion with various damage-specific endonucleases. The aDNA samples were found to be highly sensitive to alkali with an estimated frequency of >1 alkali-labile site (ALS) per 20 bp DNA. Digestion with *E. coli* AP endonuclease IV showed that about half of the ALS was associated with base loss and half "inferred to represent modified sugar residues." Further examination with *E. coli* endonuclease III indicated a high level of hydrolytic damage in the aDNA with at least one modified residue per 10 bp DNA. Taken together these data indicate that there was a very large amount of base damage in these preserved aDNA samples. Paabo found no correlation between the degree of fragmentation and the age of the specimen, although he did suggest that dessication might be an important factor in DNA degradation. Hoss et al. [46] used GC/MS to quantify several types of base damage in similarly aged aDNA samples from desiccated tissues and had somewhat different results. They were able to detect eight different types of oxidative damage in aDNA samples, with 5-OH-5-methylhydantoin, 5-OH-hydantoin, and 8-oxoguanosine predominating. They observed an inverse correlation between PCR amplification efficiency and hydantoin frequency suggesting that some types of oxidative damage may block the progression of *Taq* polymerase. Neither Paabo nor Hoss found a correlation between DNA damage and age, although Hoss did note that samples with low damage levels were primarily collected from cold environments or ice. The distribution of damage in mtDNA was indirectly analyzed be comparing sequences from a large set of cloned human aDNA specimens [47–50]. Several DNA damage hotspots manifested as A → G

and C → T transitions were evident and correlated with potential sites of deamination products of adenine and cytosine.

These observations are not consistent with laboratory studies showing that DNA damage and degradation accumulate with time (see above). It is possible (probable?) that damage in the aDNA extracted from dessicated museum specimens analyzed by Paabo and Hoss may have been at or near saturation. The DNA degradation rates determined by Lindahl and others are based on relatively short-term experiments performed at 37 °C in purified aqueous solution. Needless to say, ancient specimens do not exist under such controlled conditions and are often found at a broad range of temperatures in complex chemical environments. Paabo's suggestion that "the conditions under which the individual specimens have been preserved may be of decisive importance" may be an understatement and desiccated tissues stored at relatively high temperatures (e.g., 10–23 °C) may not be suitable substrates for examining extended rates of DNA damage and degradation. Indeed, desiccation itself is known to cause significant levels of DNA damage including strand breaks [51] and DNA-protein cross-links [52].

4. Studies on DNA damage in ancient permafrost

We began our studies on DNA damage and degradation in the permafrost samples used in the paleoecological studies discussed above [6] (Table 1). We collected 4–52 g (wet weight) samples of frozen sediment soil specimens from Siberian permafrost drill cores dated between 10.4 kyr and 1.5–2 Ma. The cores were drilled between the Lena and Kolyma rivers under carefully controlled conditions monitoring contamination from extant sources. DNA was extracted from soil specimens in a dedicated aDNA laboratory (in Copenhagen, Denmark) using previously established methods [5].

The DNA from 10.4, 19, 20–30, 300–400, and 400–600 kyr was end-labeled with ^{32}P using polynucleotide kinase (PK) or Klenow fragment. We found that the efficiency of ^{32}P incorporation was dependent on the labeling technique used and the age, integrity, and yield of the DNA substrate. Initially, DNA was end-labeled with T4 PNK, which transferred radiolabeled phosphate groups from the γ position of ATP to a 5′ OH group on DNA. Using no pretreatment, we showed incorporation of ^{32}P above background in the DNA from all three ancient sediment samples (i.e., 10.4, 390, and 500 kyr). Of particular interest, we found that the efficiency of end-labeling correlated with the age of the sample, showing significantly higher incorporation in older samples (Fig. 1), presumably due to the greater degree of fragmentation (see below). Pretreatment of the aDNA with 0.5N NaOH or shrimp alkaline phosphatase (SAP) prior to end-labeling significantly increased specific activities; this effect was compounded when both treatments were used. Because 0.5N NaOH denatures DNA and fragments DNA even further at ALS, pretreatment with SAP and subsequent mild denaturation were used for neutral and denaturing agarose gel analyses of Dsb and Ssb frequencies. Increasing the specific activity of the aDNA would serve to increase analytical sensitivity. To increase specific activity, we labeled 10–30 ng of aDNA using *E. coli* Klenow fragment and [α-^{32}P] dATP or [α-^{32}P] ddATP. Since dideoxyATP terminates strand extension, the ratio of dATP:ddATP can be adjusted to maximize ^{32}P incorporation without significant increases in molecular length. Preliminary studies show specific activities higher than PNK labeling without significant increases in fragment lengths.

Table 1
Ancient DNA sample sites and analyses performed

Year/depth (m)	Site (longitude/latitude)	Age range (B.P.)	pH	ICL	PNK	Klenow
1990/1.6	Kolyma lowland (158°28′E, 69°23′N)	Holocene 10.43 kyr	7.3	+	+	+
1993/4.0	Laptev sea coast (129°30′E, 71°40′N)	Late Pleistocene 18.98 kyr	7.9	+	ND	+
2001/4.8	Kolyma lowland (156°59′E, 69°29′N)	Late Pleistocene 20–30 kyr	7.4	+	ND	ND
2001/9.2	Lake Olyor (156°20′E, 69°25′N)	Late Pleistocene 20–30 kyr	7.5	+	ND	ND
1995/14.8	Laptev sea coast (140°10′E, 72°55′N)	Middle Pleistocene 300–400 kyr	7.3	+	+	+
1993/31.6	Khomus–Yuryakh river (153°40′E, 70°05′N)	Middle Pleistocene 400–600 kyr	7.5	+	+	+
1990/31.1	Kolyma lowland (156°59′E, 69°29′N)	Late Pliocene/Early Pleistocene 1.5–2.0 Ma	8.2	+	ND	ND

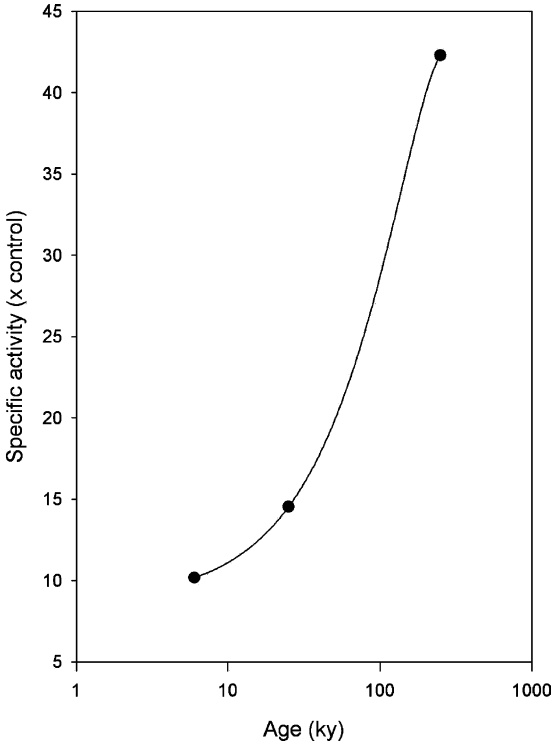

Fig. 1. DNA end-labeling afficiency and age.

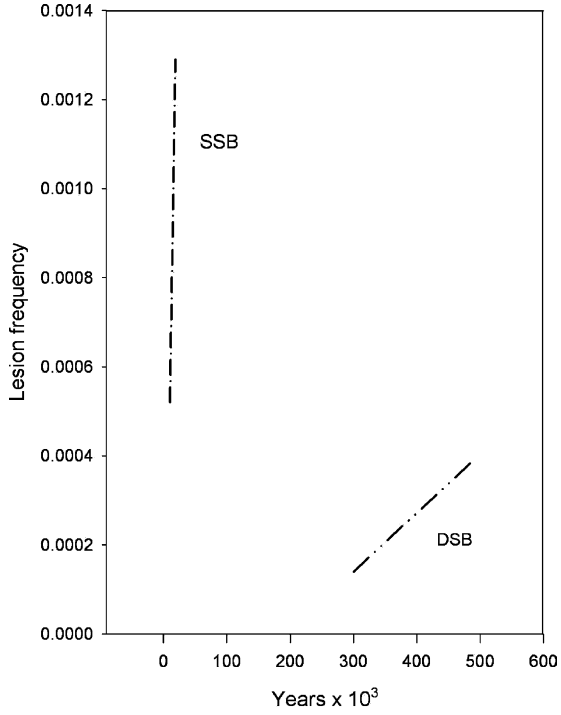

Fig. 2. Fragmentation of aDNA.

Subsequent to labeling, the aDNA was analyzed using native and denaturing gel electrophoresis. The gels were dried (O.N. at 23 °C) and exposed to a phosphorimaging plate (Molecular Dynamics) for an appropriate time consistent with the specific activity of the DNA. The plates were scanned on a phosphorimager (STORM840) and densitometry performed in ImageJ (W. Rasband, http://rsb.info.nih.gov/ij/). The dispersion function of the markers, number average molecular length (NAML) of each lane and lesion frequencies (f_{lesion}) were determined as described in Sutherland et al. [53]. The Dsb frequencies were obtained by comparing the NAML of dsDNA from 10.4 and 19 kyr and corrected for the frequency of closely opposed Ssb estimated from the Ssb determinations (Fig. 2). Rate constants were determined for Dsb and Ssb and found to be $\sim 2 \times 10^{-17}$ and $\sim 2 \times 10^{-15}$, respectively, consistent with the values estimated by Lindahl and co-workers.

We developed a very simple assay for estimating the extent of cross-links (ICLs) in DNA by examining the refractivity of samples to heat denaturation [6]. The relative yield of duplex DNA was determined using a very sensitive microplate fluorescence assay and Picogreen ©(Molecular Probes) a fluorescent DNA probe that has 10-fold greater affinity for double-stranded compared to single-stranded DNA. Hence, any loss of duplex DNA after denaturation would result in a significant reduction in fluorescence signal. This analysis assumes that inhibited denaturation results from ICLs (this is discussed in more detail below).

Comparing the extent of signal reduction in aDNA with that of an undamaged control (*E. coli*) measured at comparable concentrations the proportion of DNA refractory to denaturation was estimated (Fig. 3). Using this assay, the estimated percentage of ICLs in the DNA from the 10.4, 19, and 400–600 kyr permafrost specimens was found to follow first-order kinetics at −9 to −12 °C. Rate constants were calculated as above and shown to be $\sim 4 \times 10^{-14}$, considerably faster than Dsb or Ssb. Indeed, in very old permafrost samples (3000–4000 kyr) the DNA appeared to be completely cross-linked with the frequency of ICLs at least comparable to the frequency of strand breaks.

Because the structure of the termini associated with the fragmented aDNA may be useful in determin-

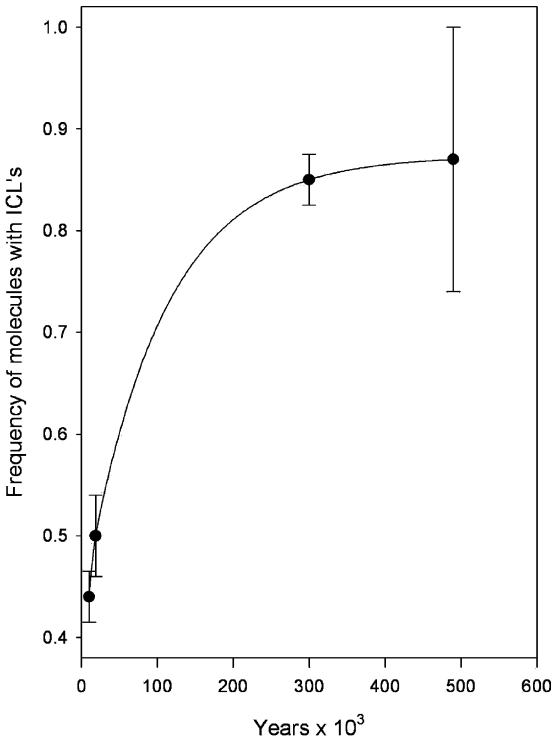

Fig. 3. Interstrand cross-link formation in aDNA.

ing the chemical or biochemical mechanisms underlying strand break formation [54], the DNA from 10.4, 400–600, and 500 kyr specimens was subjected to alkali and SAP treatments prior to PNK end-labeling (Table 2). Statistical analysis of the specific incorporation of radioactivity during the end-labeling assays revealed an overall increase with age and treatment that was different from the untreated DNA (see Fig. 1). The different structures of the 5′ termini found in aDNA from the 10.4, 300–400, and 400–600 kyr samples correspond to the end products expected by Ssb generated at AP sites, indicating that depurination or depyrimidation is most likely the main cause of Ssb formation. The relative yields of intact 5′ PO_4 groups and 5′ alkali-labile free sugar residues correspond to the pattern previously observed in depurination studies of fully hydrated DNA from phage PM2 [55]. Intact alkali-sensitive AP sites, previously observed in purified aDNA [27], could cause the increased PNK incorporation of radioactivity observed after alkali treatment. However, because the AP site degradation is a rapid and constant process [31], the amount of ^{32}P incorporated at alkali-sensitive sites should be constant over time [32]. Since this is not what we observe it is probable that, if AP-sites are present in the ancient sediment DNA, they have a minor influence on strand break frequencies.

Table 2
Effect of pretreatment on end-labeling efficiency

Treatment	10.4 kyr	390 kyr	500 kyr
Untreated	124[a]	130	217
NaOH	186	171	335
SAP	191	279	484
NaOH + SAP	241	514	456

[a] cpm × 1000 μg^{-1} DNA.

The amount of 5′ PO_4 and OH ends in the aDNA samples is approximately equivalent and is not consistent with the expected pattern of four-fold excess 5′ terminal PO_4 groups predicted by Lindahl and Andersson [55]. These data imply that processes other than depurination may be involved in the generation of Ssb and suggest that direct spontaneous cleavage of the phosphodiester backbone [56] may be responsible for the surplus 5′ OH groups. This is further supported by the presence of 3′OH ends in aDNA. Because depurination rates in vivo can range from 3×10^{-11} to 1.3×10^{-6} at 37° C [28], artefact-free levels of background damage have been difficult to determine and verify. The Ssb rate constants in the aDNA from the frozen sediments and the depurination kinetics described by Lindahl and Nyberg [28] are very similar, strongly supporting the role of spontaneous depurination in aDNA degradation. Because the specimens in the current study originated from the same tundra zone with continuous permafrost and ground temperatures between −9 and −2°C and because we found comparable rate constants at widely separated ages (not shown), the correlation between DNA damage and age most likely derives from spontaneous events unaffected by the irregular distribution of soil chemicals (e.g., metals) or any significant introduction of damage artifacts during the DNA purification process.

5. Restoration of ancient DNA

Ancient DNA contains invaluable information, pertinent to numerous fields of study including paleobiol-

ogy, anthropology, and archeology. Indeed, the science developed around aDNA studies may even impact the search for extraterrestrial life in the Martian permafrost, in the seas below the surface of the Jovian moon Europa or on Saturn's moon Titan. This information is severely restricted by the small amount of extractable DNA from ancient environments and fossils as well as the limitations imposed by extensive DNA degradation and structural damage. Even partial restoration of aDNA would enhance PCR substrate size and integrity and greatly increase the sequence information available for relevant fields of study.

Strand breaks are the predominate damage in most aDNA samples. It is very probable that Dsb cannot be repaired in aDNA. Dsb repair in the cell is contingent upon the proximity of ends held together by nuclear scaffolding or chromatin, which are lost during the DNA extraction process. Ssb's are the most dominant and ubiquitous damage associated with spontaneous degradation and at high concentrations can form Dsb. Accumulation of Ssb's can produce regions of single-stranded DNA linking double-stranded DNA regions. Such internal single-strand regions are often more vulnerable to DNA damage formation and breakage producing DNA molecules with single-strand overhangs of assorted lengths on either end of the molecule. Structures of this sort can persist indefinitely under non-metabolic conditions. Ssb's can also cause fragmentation of duplex DNA into small single-stranded molecules during the initial denaturation step of PCR amplification, greatly reducing its efficiency [57]. Considerable restoration of DNA molecules containing gapped duplexes and overhangs has been achieved using the gap-filling activity of *E. coli* DNA pol I followed by nick closure with ligase. This reconstructive polymerization step has been used successfully to increase amplification of ancient nuclear DNA from equids buried by Vesuvius at Pompeii and Herculaneum in 79 A.D. [58] as well as 7000 years-old coprolite DNA from Brazil and Chile [59].

Many different types of DNA damage are present in aDNA samples, the extent and spectrum of which are dependent on age and the environment as well as the storage and extraction conditions. Base damage has been detected in aDNA samples using *E. coli* endonucleases III and IV [27]. Using GC/MS several different types of oxidative damage in dessicated aDNA samples, including 5-OH-5-methylhydantoin, 5-OH-hydantoin, and 8-oxoguanosine have been identified [46]. Some types of base damage can either block polymerase progression, thus reducing the number of available copies for PCR analysis (e.g., hydantoin), or act as a miscoding lesion, introducing errors into the PCR product (e.g., 8-oxodeoxyguanine). Although not as yet deployed, a strategy of digesting aDNA with other repair endonucleases (e.g., FaPy glycosylase, uracil DNA glycosylase, and oxoguanine glycosylase) after reconstructive polymerization and ligation could prove useful in increasing PCR efficiency. Because some types of base damage inhibit DNA polymerase progression, the efficiency of gap-filling could be increased if such damage could be removed prior to treatment. However, base damage repair in single-stranded DNA by damage-specific endonucleases is very inefficient and would actually result in further fragmentation due to endonucleolytic scission of the single-stranded phosphate backbone.

As mentioned, in very old permafrost DNA samples the frequency of ICLs is so high that virtually every duplex DNA molecule >100–200 bp contains at least one cross-link. It is very probable that ancient desiccated remains also contain high frequencies of ICLs. ICLs represent the "double-edged sword" of ancient DNA damage: on one hand, they impose significant structural blocks to DNA amplification and their in vitro repair is problematic; on the other hand, they serve to inhibit strand separation and DNA degradation. Hence, ICLs may not only protect DNA from further degradation but may also be useful in holding DNA strands together to facilitate in vitro repair using the approaches discussed above. However, ICL repair is complex, requiring several repair systems including nucleotide excision repair (i.e., the XPF and ERCC1 proteins) to uncouple the cross-link, and homologous recombination to provide the genetic information needed to complete repair [60]. Indeed, recent work from Zhang et al. [61] implicates the mismatch repair system in ICL recognition. It is evident that, unlike other types of damage in aDNA, ICLs may prove a formidable roadblock to PCR amplification. Creative strategies have been proposed to solve this dilemma, however, and include transfection and recovery of aDNA in prokaryotic or eukaryotic hosts that are known to be either particularly resistant to cross-linking or other DNA damaging agents or particularly

robust in their repair capacity. *Deinococcus radiodurans* has been suggested as a suitable host for restoring aDNA because of its extreme radio-resistance and ability to tolerate high levels of DNA damage. Other hosts such as *Xenopus laevis* oocytes may also prove useful. A problem with transfection protocols is that they presumably will utilize host homologous recombination systems that require intact homologous host DNA to complete the ICL repair process. Ancient DNA that lacks appropriate homologous sequences may incorporate a significant amount of host sequences during the repair process.

6. Conclusions

Data from aDNA studies continue to be highly informative and significant. Our work on aDNA provides the first measured DNA degradation rates in the environment over extended periods of time and serves to test the theoretical rate constants. If these rates are fairly constant in similar environments, the extent of DNA degradation may be a useful marker for estimating sample age. Continued studies on aDNA damage will elucidate mechanisms of DNA degradation in cold environments, increase our understanding of basic DNA chemistry, and provide insight into the impact of environmental factors on these processes. Increased knowledge of the types of damage present in a given aDNA sample will facilitate interpretation of the PCR products amplified from that sample [49,50]. Indeed, increased knowledge of the types of damage and degree of fragmentation in a particular aDNA sample will help focus and facilitate the design of repair and recovery strategies. Even partial restoration of aDNA to normal structure and sequence will significantly expand the information available to PCR analysis and our understanding of ancient patterns and processes.

Acknowledgments

We thank Becky Brooks for assistance in manuscript preparation and Lakshmi Paniker for technical assistance. This work was supported in part by grants from the National Science Foundation OPP-0085589 and NIEHS Center Grant ES07784.

References

[1] T.A. Brown, K.A. Brown, Ancient DNA: using molecular biology to explore the past, Bioessays 16 (1994) 719–726.
[2] A. Cooper, A. Rambaut, V. Macaulay, E. Willerslev, A.J. Hansen, C. Stringer, Human origins and ancient human DNA, Science 292 (2001) 1655–1656.
[3] M. Hofreiter, D. Serre, H.N. Poinar, M. Kuch, S. Paabo, Ancient DNA, Nat. Rev. Genet. 2 (2001) 353–359.
[4] F.A. Kaestle, K.A. Horsburgh, Ancient DNA in anthropology: methods, applications, and ethics, Am. J. Phys. Anthropol. Suppl. 35 (2002) 92–130.
[5] E. Willerslev, A.J. Hansen, B. Christensen, J.P. Steffensen, P. Arctander, Diversity of Holocene life forms in fossil glacier ice, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 8017–8021.
[6] E. Willerslev, A.J. Hansen, J. Binladen, T.B. Brand, M.T. Gilbert, B. Shapiro, M. Bunce, C. Wiuf, D.A. Gilichinsky, A. Cooper, Diverse plant and animal genetic records from Holocene and Pleistocene sediments, Science 300 (2003) 791–795.
[7] R. Higuchi, B. Bowman, M. Freiberger, O.A. Ryder, A.C. Wilson, DNA sequences from the quagga, an extinct member of the horse family, Nature 312 (1984) 282–284.
[8] C. Hanni, V. Laudet, D. Stehelin, P. Taberlet, Tracking the origins of the cave bear (*Ursus spelaeus*) by mitochondrial DNA sequencing, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 12336–12340.
[9] M. Hoss, A. Dilling, A. Currant, S. Paabo, Molecular phylogeny of the extinct ground sloth Mylodon darwinii, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 181–185.
[10] R.H. Thomas, W. Schaffner, A.C. Wilson, S. Paabo, DNA phylogeny of the extinct marsupial wolf, Nature 340 (1989) 465–467.
[11] D.N. Janczewski, N. Yuhki, D.A. Gilbert, G.T. Jefferson, S.J. O'Brien, Molecular phylogenetic inference from saber-toothed cat fossils of Rancho La Brea, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 9769–9773.
[12] M. Noro, R. Masuda, I.A. Dubrovo, M.C. Yoshida, M. Kato, Molecular phylogenetic inference of the woolly mammoth Mammuthus primigenius, based on complete sequences of mitochondrial cytochrome b and 12S ribosomal RNA genes, J. Mol. Evol. 46 (1998) 314–326.
[13] A. Cooper, C. Lalueza-Fox, S. Anderson, A. Rambaut, J. Austin, R. Ward, Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution, Nature 409 (2001) 704–707.
[14] D.M. Lambert, P.A. Ritchie, C.D. Millar, B. Holland, A.J. Drummond, C. Baroni, Rates of evolution in ancient DNA from Adelie penguins, Science 295 (2002) 2270–2273.
[15] M. Hofreiter, H.N. Poinar, W.G. Spaulding, K. Bauer, P.S. Martin, G. Possnert, S. Paabo, A molecular analysis of ground sloth diet through the last glaciation, Mol. Ecol. 9 (2000) 1975–1984.
[16] M. Scholz, L. Bachmann, G.J. Nicholson, J. Bachmann, I. Giddings, B. Ruschoff-Thale, A. Czarnetzki, C.M. Pusch, Genomic differentiation of Neanderthals and anatomically modern man allows a fossil-DNA-based classification of morphologically indistinguishable hominid bones, Am. J. Hum. Genet. 66 (2000) 1927–1932.

[17] I.V. Ovchinnikov, A. Gotherstrom, G.P. Romanova, V.M. Kharitonov, K. Liden, W. Goodwin, Molecular analysis of Neanderthal DNA from the Northern Caucasus, Nature 404 (2000) 490–493.
[18] P. Endicott, M.T. Gilbert, C. Stringer, C. Lalueza-Fox, E. Willerslev, A.J. Hansen, A. Cooper, The genetic origins of the Andaman Islanders, Am. J. Hum. Genet. 72 (2003) 178–184.
[19] S.W. Carlyle, R.L. Parr, M.G. Hayes, D.H. O'Rourke, Context of maternal lineages in the Greater Southwest, Am. J. Phys. Anthropol. 113 (2000) 85–101.
[20] A.R. Zink, U. Reischl, H. Wolf, A.G. Nerlich, Molecular analysis of ancient microbial infections, FEMS Microbiol. Lett. 213 (2002) 141–147.
[21] A.R. Zink, W. Grabner, U. Reischl, H. Wolf, A.G. Nerlich, Molecular study on human tuberculosis in three geographically distinct and time delineated populations from ancient Egypt, Epidemiol. Infect. 130 (2003) 239–249.
[22] N. Konomi, E. Lebwohl, K. Mowbray, I. Tattersall, D. Zhang, Detection of mycobacterial DNA in Andean mummies, J. Clin. Microbiol. 40 (2002) 4738–4740.
[23] A.C. Aufderheide, W. Salo, M. Madden, J. Streitz, J. Buikstra, F. Guhl, B. Arriaza, C. Renier, L.E. Wittmers Jr., G. Fornaciari, M. Allison, A 9000-year record of Chagas' disease, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 2034–2039.
[24] H.N. Poinar, M. Kuch, K.D. Sobolik, I. Barnes, A.B. Stankiewicz, T. Kuder, W.G. Spaulding, V.M. Bryant, A. Cooper, S. Paabo, A molecular analysis of dietary diversity for three archaic native Americans, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 4317–4322.
[25] A. Sidow, A.C. Wilson, S. Paabo, Bacterial DNA in Clarkia fossils, Philos. Trans. R. Soc. Lond. B Biol. Sci. 333 (1991) 432–433 (Discussion 429–432).
[26] H. Poinar, M. Kuch, G. McDonald, P. Martin, S. Paabo, Nuclear gene sequences from a late pleistocene sloth coprolite, Curr. Biol. 13 (2003) 1150–1152.
[27] S. Paabo, Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification, Proc. Natl. Acad. Sci. U.S.A. 86 (1989) 1939–1943.
[28] T. Lindahl, B. Nyberg, Rate of depurination of native deoxyribonucleic acid, Biochemistry 11 (1972) 3610–3618.
[29] T. Lindahl, O. Karlstrom, Heat-induced depyrimidination of deoxyribonucleic acid in neutral solution, Biochemistry 12 (1973) 5151–5154.
[30] L.A. Loeb, B.D. Preston, Mutagenesis by apurinic/apyrimidinic sites, Annu. Rev. Genet. 20 (1986) 201–230.
[31] T. Lindahl, Instability and decay of the primary structure of DNA, Nature 362 (1993) 709–715.
[32] R. Shapiro, Damage to DNA caused by hydrolysis, in: E. Seeberg, K. Kleppe (Eds.), Chromosome Damage and Repair, Plenum Press, New York, 1981.
[33] T. Lindahl, B. Nyberg, Heat-induced deamination of cytosine residues in deoxyribonucleic acid, Biochemistry 13 (1974) 3405–3410.
[34] T. Lindahl, DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair, Prog. Nucleic Acid Res. Mol. Biol. 22 (1979) 135–192.
[35] M. Ehrlich, X.Y. Zhang, N.M. Inamdar, Spontaneous deamination of cytosine and 5-methylcytosine residues in DNA and replacement of 5-methylcytosine residues with cytosine residues, Mutat. Res. 238 (1990) 277–286.
[36] C.I. Smith, A.T. Chamberlain, M.S. Riley, C. Stringer, M.J. Collins, The thermal history of human fossils and the likelihood of successful DNA amplification, J. Hum. E 45 (2003) 203–217.
[37] S.J. Dancer, P. Shears, D.J. Platt, Isolation and characterization of coliforms from glacial ice and water in Canada's high Arctic, J. Appl. Microbiol. 82 (1997) 597–609.
[38] S.S. Abyzov, N.E. Bobin, B.B. Koudriashov, Microbiological flora as a function of ice depth in central Antarctica, Life Sci. Space Res. 17 (1979) 99–103.
[39] S.S. Abyzov, N.F. Kirillova, G.V. Cherkesova, et al., Prolonged anabiosis for sporogenous bacteria in the interior of the Central Antarctic Glacier, Akademiia nauk SSSR, Biol. Bull. 15 (1989).
[40] S.S. Abyzov, V.Y. Lipenkov, N.E. Bobin, B.B. Kudriashov, V.M. Pashkevich, Characteristics of different strata of the central glacier in connection with microbiological studies, Antarktika 32 (1993).
[41] S.S. Abyzov, V.I. Biriuzova, N.A. Kostrokina, et al., Actinomycete from very ancient layers of the icesheet of central Antarctica, Microbiology 59 (1990) 768–773 (Translated from Mikrobiologiia).
[42] S.S. Abyzov, I.N. Mitskevich, M.N. Poglazova, N.I. Barkov, V.Y. Lipenkov, N.E. Bobin, B.B. Koudryashov, V.M. Pashkevich, M.V. Ivanov, Microflora in the basal strata at Antarctic ice core above the Vostok lake, Adv. Space Res. 28 (2001) 701–706.
[43] T.A. Sorokina, S.S. Abyzov, et al., Pseudomonas bacteria from an antarctic glacial ice sheet Akademiia nauk SSSR. Izvestiya., Biol. Bull. 13 (1986).
[44] D.A. Gilichinsky, E.A. Vorobyova, L.G. Erokhina, D.G. Fyordorov-Davydov, N.R. Chaikovskaya, D.G. Fyordorov-Dayvdov, Long-term preservation of microbial ecosystems in permafrost, Adv. Space Res. 12 (1992) 255–263.
[45] J. Zhou, M.E. Davey, J.B. Figueras, E. Rivkina, D. Gilichinsky, J.M. Tiedje, Phylogenetic diversity of a bacterial community determined from Siberian tundra soil DNA, Microbiology 143 (Pt 12) (1997) 3913–3919.
[46] M. Hoss, P. Jaruga, T.H. Zastawny, M. Dizdaroglu, S. Paabo, DNA damage and DNA sequence retrieval from ancient tissues, Nucleic Acids Res. 24 (1996) 1304–1307.
[47] M. Hofreiter, V. Jaenicke, D. Serre, A. Haeseler Av, S. Paabo, DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA, Nucleic Acids Res. 29 (2001) 4793–4799.
[48] A. Hansen, E. Willerslev, C. Wiuf, T. Mourier, P. Arctander, Statistical evidence for miscoding lesions in ancient DNA templates, Mol. Biol. E 18 (2001) 262–265.
[49] M.T. Gilbert, A.J. Hansen, E. Willerslev, L. Rudbeck, I. Barnes, N. Lynnerup, A. Cooper, Characterization of genetic miscoding lesions caused by postmortem damage, Am. J. Hum. Genet. 72 (2003) 48–61.
[50] M.T. Gilbert, E. Willerslev, A.J. Hansen, I. Barnes, L. Rudbeck, N. Lynnerup, A. Cooper, Distribution patterns of postmortem

damage in human mitochondrial DNA, Am. J. Hum. Genet. 72 (2003) 32–47.
[51] K. Dose, A. Bieger-Dose, M. Labusch, M. Gill, Survival in extreme dryness and DNA-single-strand breaks, Adv. Space Res. 12 (1992) 221–229.
[52] A. Bieger-Dose, K. Dose, R. Meffert, M. Mehler, S. Risi, Extreme dryness and DNA-protein cross-links, Adv. Space Res. 12 (1992) 265–270.
[53] B.M. Sutherland, P.V. Bennett, J.C. Sutherland, DNA damage quantitation by alkaline gel electrophoresis, Methods Mol. Biol. 113 (1999) 183–202.
[54] W.K. Pogozelski, T.D. Tullius, Oxidative strand scission of nucleic acids: routes initiated by hydrogen abstraction from the sugar moiety, Chem. Rev. 98 (1998) 1089–1108.
[55] T. Lindahl, A. Andersson, Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid, Biochemistry 11 (1972) 3618–3623.
[56] P. Crine, W.G. Verly, Determination of single-strand breaks in DNA using neutral sucrose gradients, Anal. Biochem. 75 (1976) 583–595.
[57] C.M. Pusch, I. Giddings, M. Scholz, Repair of degraded duplex DNA from prehistoric samples using *Escherichia coli* DNA polymerase I and T4 DNA ligase, Nucleic Acids Res. 26 (1998) 857–859.
[58] G. Di Bernardo, S. Del Gaudio, M. Cammarota, U. Galderisi, A. Cascino, M. Cipollaro, Enzymatic repair of selected cross-linked homoduplex molecules enhances nuclear gene rescue from Pompeii and Herculaneum remains, Nucleic Acids Res. 30 (2002) 16.
[59] A.M. Iniguez, A. Araujo, L.F. Ferreira, A.C. Vicente, Analysis of ancient DNA from coprolites: a perspective with random amplified polymorphic DNA-polymerase chain reaction approach, Mem. Inst. Oswaldo Cruz 98 (Suppl. 1) (2003) 63–65.
[60] P.J. McHugh, V.J. Spanswick, J.A. Hartley, Repair of DNA interstrand cross-links: molecular mechanisms and clinical relevance, Lancet Oncol. 2 (2001) 483–490.
[61] N. Zhang, X. Lu, X. Zhang, C.A. Peterson, R.J. Legerski, hMutSbeta is required for the recognition and uncoupling of psoralen interstrand cross-links in vitro, Mol. Cell Biol. 22 (2002) 2388–2397.