Acc. Chem. Res. **2002**, *35*, 676−684

The Genetic Secrets Some Fossils Hold

HENDRIK N. POINAR*
Max Planck Institute, Evolutionary Anthropology, Inselstrasse 22, D-04103 Leipzig, Germany

Received June 26, 2001

ABSTRACT
Most animals that once lived have gone extinct. The remains of a few of these can be found in museum collections worldwide. As modern evolutionary biology is limited to the use of extant taxa, retrieving DNA from extinct or subfossil organisms can add significant insight into past population history and resolve phylogenies that can be tentative by morphology alone. DNA is a relatively weak molecule, comparatively speaking, yet under certain conditions it persists in the fossil record, despite what in vitro chemistry predicts. While most fossil remains do not contain DNA, museum specimens can be screened for the presence of conditions that would be conducive for nucleic acid preservation by measuring the extent of amino acid racemization and by looking at the extent of protein hydrolysis by pyrolysis gas chromatography/mass spectrometry. Results from these types of analyses suggest that the preservation of DNA is linked to the temperature and its constancy at a site rather than its age. Chemical analyses of coprolites from extinct herbivores from the late Pleistocene, as well as Archaic Native Americans, show the presence of compounds from the Maillard reaction. Upon the cleaving of these products, the defecator can be identified and his diet analyzed.

Introduction

Over the course of geological time, some 50 billion species have lived on our planet. Today between 4 and 40 million species currently inhabit our earth, which means that roughly 99.9% of all organisms that once existed have gone extinct.[1] The remains of a small percentage of those extinct animals or the trace fossils that they left behind exist in a few token museum collections worldwide. It has been my conviction to make these fossils samples amenable to current DNA techniques, in the hopes of resolving evolutionary and ecological questions that are otherwise impossible with current morphological techniques. Most evolutionary biologists are limited to the use of modern taxa to resolve phylogenies and are thus "time trapped". Access to ancient specimens would allow the evolutionary biologist a true look into the genetic past of extinct animals.

The first ancient DNA reports stemmed from the remains of recently extinguished species, namely the horse-like Quagga and the marsupial wolf,[2,3] both only about 100 years old from museum collections. It was the development, shortly thereafter, of a novel technique known as the polymerase chain reaction (PCR) which opened the field to the study of many more fossil remains.[4] PCR is a unique enzymatic process that involves the repeated copying of a targeted stretch of DNA many times during a 3-h repetitive cycling process. It is unique in that it is theoretically possible to target a single copy and amplify it a million-fold, so that it can be cloned (placed into bacteria) and sequenced. No other method of detection, as far as I know, is as sensitive. Alas, this sensitivity is also PCR's downfall, as traces of contaminating DNA from contemporary sources, such as museum curators, pose a big threat to the field, especially for ancient human DNA analysis, where it may be particularly difficult to weed out the exogenous from the endogenous DNA sequences.[5,6] Many authors have repeatedly put forth a plethora of "criteria for authenticity" that one should follow. It is not in the scope of this Account to address these particular issues, and the reader is referred to several reviews on the subject.[7−10]

Despite the contamination issues, there have been many successes in the field in recent years which have added to our understanding of various aspects of modern biology. For example, the recent extraction and sequencing of a small portion of mitochondrial DNA from three Neandertals[11,12] shows them to fall clearly outside the variation of modern humans. Thus it is unlikely, although not impossible, that Neandertals interbred with anatomically modern humans, despite several years of overlap in certain parts of Western Asia.

Before one can "open" fossils and make then amenable to current DNA technology, one must first understand the limitations of the DNA molecule itself: that is, how stable is it, under what conditions will it preserve, and what are the types of reactions that will limit its survival in the fossil record? In the following paragraphs I briefly describe the major routes of DNA degradation as learned from both in vitro chemical degradation and past DNA analyses from fossil remains.

DNA Degradation and Preservation

Unfortunately for the paleogeneticist, comparatively speaking, the DNA molecule is one of the least stable molecules within our cells (Table 1). It is the subject of rapid hydrolysis and oxidation over short time periods, limiting its "life" in vivo and within the environment.[7]

Once an organism dies, presumably the single most important factor in the long-term preservation of its DNA is the rate at which specific cellular enzymes, called nucleases, can be stopped. Apoptosis results in the efficient and rapid cleavage of DNA into small fragments. However, as these are energy-requiring functions, and a cell without oxygen will deplete its energy sources quickly, nuclease degradation may cease relatively soon post mortem. The organism must then face the bacterial, fungal, and insect onslaught which can be quite effective

Hendrik N. Poinar was born in Utrecht, Holland, in 1969. He received his B.S. in biochemistry in 1992 and his M.S. in biology in 1994 from the California Polytechnic State University in San Luis Obispo, CA, and his Ph.D. from Ludwig Maximillians Universität München in 1999. He is currently a small group leader at the newly formed Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, where his group focuses on the preservation and degradation of nucleic acids in fossil specimens.

* E-mail: poinar@eva.mpg.de. Phone: 49 341−9952 508. Fax: 49 341 9952-555.

10.1021/ar000207x CCC: $22.00 © 2002 American Chemical Society
Published on Web 02/02/2002

Table 1. Cellular Compounds, Their Susceptible Bonds and Groups, and Their Preservation Potential

compounds class	susceptible bonds	susceptible groups	preservation potential[a]
DNA, RNA	phosphate esters, glycosidic bond, 5,6 C–C bond pyrimidines	heterocyclic rings, amino groups, methyl group	–
proteins	peptide bond	side chain, chiral center	–/+
carbohydrates	acetal	hydroxy, amide	+
lipids	ester, ether, amide	hydroxyl, carboxyl, ester	++
cutin	ester, ether	hydroxyl, carboxyl	+++
lignin	ether	methoxyl aromatic rings	++++

[a] – to ++++ (weakest to strongest), a rough estimate of the preservation potential for an unaltered molecule based upon relative bond strengths.

but often incomplete.[13] Once bacterial onslaught has slowed, the DNA molecule is still subject to chemical degradation via hydrolysis and oxidation. To understand the processes which degrade the DNA molecule in the fossil record and under what conditions these reactions are minimized, one needs to look briefly at the molecule itself and its susceptible sites.

DNA Molecule

The DNA molecule is a polymer composed of 2′-deoxyribose sugar units linked to each other via phosphate ester links.[14] To every sugar, a purine (adenine or guanine) or a pyrimidine (cytosine or thymine) is attached at the 1′ position by a glycosidic bond. In vivo, the DNA molecule is supported by water molecules in the major grooves, thus the molecule is never really dry. In addition, nuclear DNA (not mitochondrial DNA) is wound tightly around histone proteins which presumably absorb some of the surrounding damage. Despite this, the DNA molecule is labile and prone to many forms of damage, all of which presumably limit the molecule's "half-life" in the geosphere.

Hydrolytic DNA Damage

The DNA molecule is particularly prone to hydrolytic damage[7,15] (Figure 1). The removal of the ribose sugar's 2′-OH group, creating the deoxyribose sugar, does afford the bonds joining two nucleotides (phosphodiester bond) increased strength; however, it weakens the glycosidic bond joining the bases to the sugars. Diesters, like the bonds in the phosphate sugar backbone, are normally quite labile and subject to quick hydrolytic cleavage. It has been estimated that direct cleavage of the phosphate backbone is probably the most frequent type of hydrolytic damage the DNA molecule must cope with, generating single-stranded nicks. In vivo, in a fully hydrated system this event takes place about once every 2.5 h.[15] It has been estimated that under dry conditions this rate drops some 20-fold.[16]

The glycosidic bond is also subject to direct hydrolytic attack. Base protonation, making the base a better leaving group, will cause the cleavage of the glycosidic bond, termed depurination, and form what is known as a

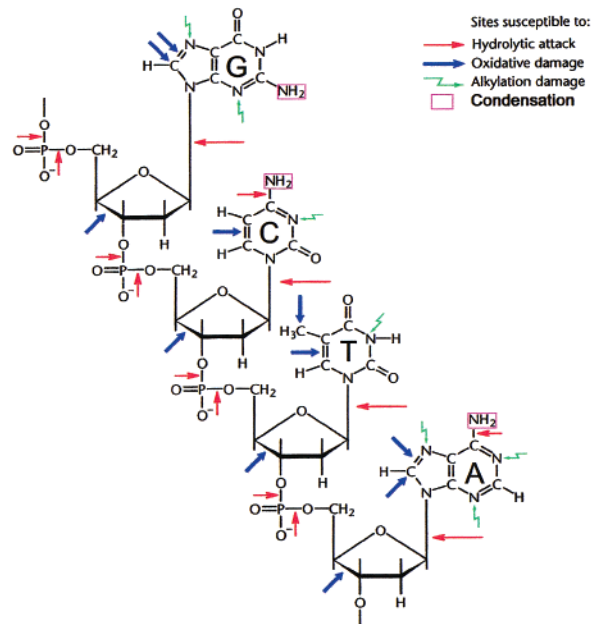

FIGURE 1. A four-base-pair stretch of single-stranded DNA indicating sites of hydrolytic, oxidative, methylation, ethylation, and possible condensation type reactions. Adenine (A), cytosine (C), guanine (G), and thymine (T).

baseless or an apurinic site (AP site). Estimated rates of depurination in double- stranded DNA suggest one event takes place about every 10 h in vivo. The rate of depurination is increased and dependent upon the temperature, the ionic strength, the pH, and heavy metal ion chelation.

Once a nucleotide base is released from DNA, the AP site can undergo cleavage, termed β-elimination, and thus result in a single-stranded nick. AP site hydrolysis occurs at a rate similar to or slightly slower than depurination and probably takes place a few days after depurination. At these rates of hydrolytic damage, a small DNA fragment of a few hundred base pairs would not be expected to survive beyond 10^4 years in most temperate settings, 10^5 years in colder ones such as the permafrost, and perhaps even 10^6 years in exceptional cases such as fossil glacial ice cores.[17] The fact that most DNA retrieved from fossil or even subfossil samples is consistently around 100–500 base pairs (bp) in length would suggest that hydrolytic DNA damage takes place relatively soon post mortem and thus is one of the more important pathways reducing its life in the geosphere.

Apart from the hydrolytic cleavage of the phosphodiester and the glycosidic bonds, DNA bases with secondary amino groups such as adenine, cytosine, 5-methylcytosine, and guanine can undergo deamination, the hydrolytic cleavage of their amino groups, resulting in hypoxanthine, uracil, thymine, and xanthine, respectively. Deamination has recently been shown to be a prominent component of some fossil DNA remains.[10]

Oxidative DNA Damage

Even when the remains of an organism "dry" (although there is nothing in nature that is 100% dry), the molecular

components of a cell are still susceptible to oxidative attack. Free radicals such as peroxide radicals ($^{\bullet}O_2$), hydrogen peroxide (H_2O_2), and hydroxy radicals ($^{\bullet}OH$) are all endogenously occurring products and hence are an important source of endogenous DNA damage[7,18] and are likely to play an important role in limiting the life of DNA in the fossil record. They may also derive from exogenous sources such as ionizing radiation, UV light, and cellular processes during bacterial and fungal degradation. Recently DNA extractions from fossil remains were subject to gas chromotagraphy/mass spectrometry in an attempt to identify oxidative base damage.[19] From all samples where no endogenous DNA could be retrieved, higher levels of two oxidative forms of base damage, 5-hydroxyhydantoin and 5-hydroxy-5-methylhydantoin, were detected. As these and many other oxidative lesions block the polymerase in PCR, and hence its ability to make copies, oxidative damage will also limit the successful retrieval of DNA from fossil remains.

Both hydrolysis and oxidation limit the survival of DNA in the fossil record, and it is likely for these reasons that few fossil samples still contain authentic endogenous DNA. As all DNA extraction methods to date are destructive methods and can require as much as a few grams of material, most museum curators are reasonably hesitant to release prized possessions, knowing that there is a small chance that something may have survived in the specimen. For these reasons it is important to screen samples prior to their extraction, to assess the state of molecular preservation.

Screening Samples for the Preservation of DNA

The methodology used for screening samples should assess the preservation of a major molecular organic component of the sample which has bond strength relatively stronger than or equal strength to that in DNA. In addition, the method should require minimal amounts of a sample (ca. 5−10 mg), be easy to measure and analyze, and be relatively quick and inexpensive to use. Finally, the system should be available for anthropologists, paleontologists, and molecular biologists alike. Assessing the level of protein preservation in fossil samples seems well suited for this task. Proteins form the bulk of most fossilized tissues such as bone, skin and teeth, are usually in high concentration per gram of tissue, and can be easily analyzed by many current techniques.

We have explored two possible avenues along these lines. One method consists of measuring the extent of amino acid racemization along with total amino acid content through fluorescent derivitization of the amino acids and separation and detection with high-pressure liquid chromatography (HPLC) and a fluorescence detector. The other method involves measuring the ratio of single amino acids (in amounts) to dipeptide moieties still present in a fossil specimen by pyrolysis gas chromatography/mass spectrometry (Py-GC/MS).

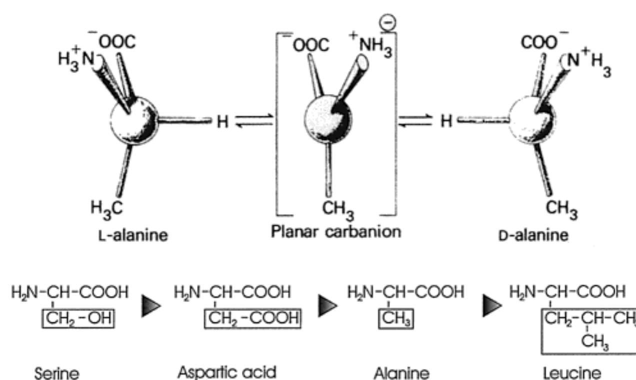

FIGURE 2. (a) Mechanism of amino acid racemization for alanine with planar carbanion intermediate shown. Re-addition of the proton occurs in water. (Redrawn from ref 20.) (b) Amino acid structures for serine, aspartic acid, alanine, and leucine given in increasing order of their electronegativities.

Amino Acid Racemization

Almost all amino acids can occur in two optical isomers, termed L and D. Both enantiomers are identical mirror images of each other and rotate plane-polarized light in opposite directions. Most naturally occurring proteins contain only L-amino acids. Metabolically active tissues contain specialized enzymes known as racemases which maintain a disequilibrium in our cells of only the L isomers.

Once an organisms dies and its enzymes cease to function, amino acids undergo racemization, that is, the interconversion of L-amino acids to D-amino acids through abstraction of the α proton, leading to a planar carbanion intermediate and re-addition of a proton to either side of the intermediate[20] (Figure 2a). This reaction observes reversible first-order kinetics. The rate of racemization of amino acids depends predominantly upon the electron-withdrawing capabilities of the R substituents: the more electron-withdrawing the R group, the faster the rate of racemization. Thus, the relative rates of racemization of free amino acids can be predicted according to their R groups. The rate of racemization of free amino acids has been measured in vitro and follows the order predicted by the electron-withdrawing capabilities of their R groups.[20] Thus, the rate of racemization of serine is faster than that of aspartic acid, which is faster than that of alanine, which in turn is faster than that of leucine[21] (Figure 2b). The rate of racemization is also increased at elevated temperatures in acidic and basic conditions and by heavy metal ion chelation, thus many of the same conditions affecting the rate of DNA depurination and thus strand cleavage.

Aspartic acid is one of the fastest racemizing amino acids and has been used in the past to date certain fossil remains. While this application is effective in a closed system, with a constant known temperature, almost all fossilized remains such as bones and teeth are open systems; they are a mixture of bacterial, fungal remnants, various minerals, and other elements and have very complicated temperature histories. Despite this, one can assess the "authenticity" of a fossil's amino acid content by looking at the relative amounts of racemization of more

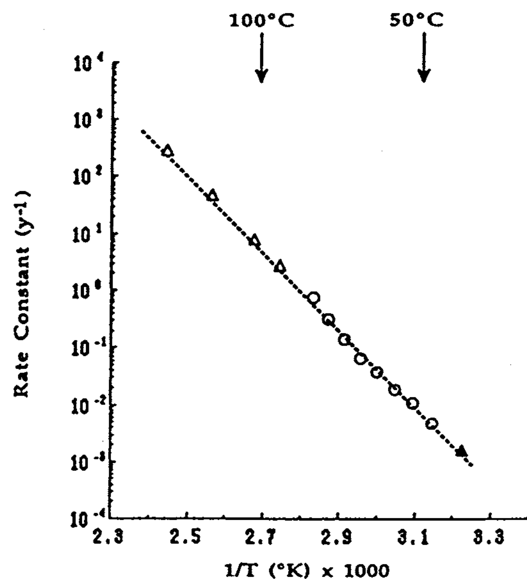

FIGURE 3. Rate of depurination of native DNA and rate of aspartic acid racemization as a funtion of 1/*K* (K) at physiological pH: △, measured rate of aspartic acid racemization; ○, rate of DNA depurination; ▲, in vivo rate of aspartic acid racemization. (Reprinted with permission from ref 22. Copyright 1994 Elsevier Science.)

than a single amino acid. When modern bone is heated in a closed system, the relative rates of racemization exhibit the predicted hierarchy as seen with free amino acids. In a fossil, where amino acids from exogenous sources are integrated into fossil bones, the ratios of those amino acids will become skewed as amino acids in different relative concentrations than those typical of bone will be added to the system. Thus, when the endogenous protein content of a fossil remain has dropped significantly through hydrolysis and leaching, a heavily contaminated system will show a D/L-aspartic acid ratio less than the D/L-alanine ratio and typically a very different amino acid distribution as is standard of most collagen-containing tissues. This is a good indication of recent contamination in a fossil system.

The rate of aspartic acid racemization is nearly identical to the rate of DNA depurination, one of the most important reactions likely involved in limiting the half-life of DNA in the geosphere.[22] In fact, both rates are almost identical over a wide temperature range at physiological pH (Figure 3). In addition, as mentioned earlier, many of the factors which increase the rate of amino acid racemization also increase the rate of DNA depurination. Thus, theoretically it should be possible to use the extent of aspartic acid racemization as an indicator of the state of chemical preservation in a fossil sample and in turn indicate whether a fossil is likely to contain endogenous nucleic acids.

To test the possible correlation of these two distinct reactions, we screened a sample set of 26 fossil remains, some of which had reproducibly yielded endogenous DNA and others which had consistently failed to yield any DNA. In all cases where the ratio of D- to L-aspartic acid was lower than 0.10 yet greater than the ratio of D/L-ala and D/L-leu, DNA was present. Conversely, when the level of aspartic acid racemization was greater than 0.15 or lower than the ratio of D/L-ala or D/L-leu, no endogenous DNA was retrievable. In addition, in the eight samples which yielded endogenous DNA, the total concentration (in parts per million) of the eight amino acids which were measured was between 90 and 100% of what a modern bone contains. Finally, the distribution in mole percent of those eight amino acids followed a predicted distribution similar to that seen in modern bone.[17]

Thus, by measuring the extent of racemization and by looking at the amino acid content, it is possible to screen samples for conditions which appear chemically promising for the preservation of DNA. As HPLC with fluorescent detection can identify amino acids at femptogram levels, one can successfully screen samples with ca. 5 mg. With an automated system one can screen about 25 samples a day and thus screen from a large museum collection, selecting only those few samples that appear chemically promising without doing more damage to the samples than is necessary.

Pyrolysis Gas Chromatography/Mass Spectrometry

Another method we approached for screening samples was the use of pyrolysis coupled with gas chromatography/mass spectrometry (Py-GC/MS). This method has been used to characterize otherwise intractable organic molecules that derive from more complex materials such as lignocellulose, chitin, resins, and plant cuticles.[23]

Flash pyrolysis involves the rapid heating of a sample on platinum wire to ca. 650 °C. This heat breaks all components to individual "monomers" which are separated according to mass and charge in the gas chromatograph and then mass counted by the spectrometer. These individual components can then be ascribed a chemical structure, and the original "polymeric" puzzle can be pieced back together. To assess its efficacy as a method of screening samples, we pyrolyzed 11 fossil samples, again some of which had reproducibly yielded DNA and others which continuously failed.[24]

Collagen is the most abundant protein in mammalian skin and bone and consists of three polypeptide chains of which glycine (ca. 33%), proline (12%), alanine (10%), and hydroxyproline (ca. 10%) are the major constituents. Once pyrolyzed, collagen produces a known set of single amino acids, as well as a set of what are known as 2,5-diketopiperazines (DKPs). Diketopiperazines are the result of a dipeptide rearrangement, known as the McLafferty rearrangement, formed during the pyrolysis of dipeptides when proline is one of the two amino acids. In collagen and bone, the most significant 2,5-DKPs formed during pyrolysis are from the following dipeptide sequences: Pro-Ala, Pro-Gly, Pro-Arg, Pro-Hyp, and Pro-Pro.

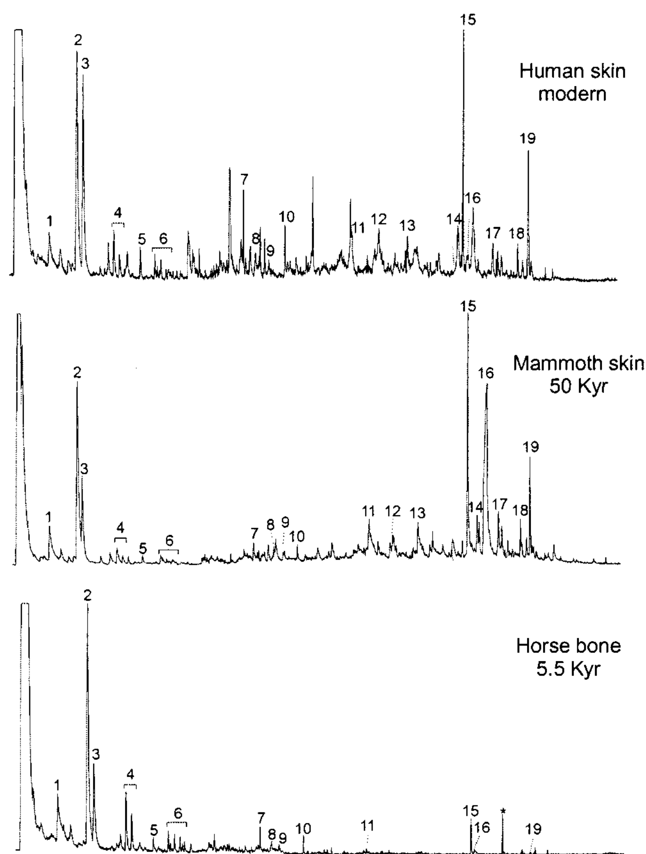

FIGURE 4. Total ion chromatograms of modern skin tissue, a 50 000-year-old mammoth tissue from the Siberian permafrost, and a 5000-year-old horse bone from a temperate site in southern Germany. Peaks numbered 1–13 are products of single amino acids; peaks numbered 14–19 are dipeptide rearrangements known as diketopiperazines (DKP).

Flash pyrolysis of 11 fossil samples (except for two where no organic residues remain) yielded a complex set of compounds. Most of these peaks are attributable to single amino acids or the 2,5-DKPs seen in the pyrolysis of modern skin tissue or purified collagen. As seen in Figure 4, modern skin tissue, which consists mainly of collagen, consists of a complex set of compounds. These are mostly attributable to derivatives of proline, glycine, alanine, and hydroxyproline. While there was no clear correlation between the relative abundance of single amino acids and the preservation of endogenous DNA, the relative abundance of the 2,5-DKP does show a correlation with DNA. The ratio of total peak height, summed for all single amino acids to the total peak height of all DKPs (AA/DKP), revealed a distinct hierarchy of preservation. It is likely that this ratio simply measures the extent of peptide hydrolysis of the intact protein within the sample. Thus, a 50 000-year-old mammoth tissue from the permafrost of Siberia is far better preserved than a 5000-year-old horse from the temperate regions of central Europe preserved in sediment (Figure 4).

The Importance of Screening Samples

Both methodologies, measuring the extent of amino acid racemization, and measuring the extent of peptide hydrolysis through pyrolysis GC/MS, represent novel ways to screen samples for conditions which would suggest the presence of endogenous DNA. Both methods require trace sample amounts of material, measure the bulk organic material (proteins) which make up bone and mammalian skin, are sensitive methods, and are readily automatable.

Despite a clear empirical correlation between the preservation of proteins and DNA in fossil samples, both methods measure only the extent of hydrolytic damage to proteinaceous material. Hydrolysis is only one chemical reaction type involved in degradation of organic matter in the geosphere, albeit probably a particularily important one, namely for DNA preservation. However, these types of analyses do not take into consideration other important reactions that are likley to take place in the degradation of fossil material, namely oxidation and condensation type reactions. It will be important in the future to use additional methods of analysis, which measure the extent of protein oxidation or perhaps of lipid or fatty acid oxidation, to shed more light on the amount of oxidative damage a fossil has seen over the course of its burial history.

What is quite clearly revealed from both data sets is that the preservation of DNA has no general correlation with the age of a fossil, but rather with the burial temperature and environmental setting. The permafrost remains of woolly mammoths from Siberia, which date to older than 50 000 years, are clearly better preserved than the temperate remains of animals only a few thousand years in age or even a 50-year-old horse from the Mojave Desert of California. Thus, in general, the colder an environment, the better the long-term preservation of the DNA. Reaction rates in general are lowered 3–4 times for every 10 °C drop in temperature.

Molecular Ecology of the Extinct Shasta Ground Sloth *Nothrotheriops shastensis*

Several caves in the American southwest contain the organic remains of extinct large herbivores that once roamed the continent. These trace fossils are mainly represented as the desiccated remains of feces, known as coprolites. These feces are in fact so extensive in some cave localities that they are referred to as dung blankets (Figure 5). While caves are known to be a great source of fossilized remains, such as bones and teeth, it is remarkable to think of the preservation of an "open" source, such as feces are, as preservable under most geological settings. These individual coprolites (Figure 6) appear morphologically as desiccated horse manure and to the layperson only a few years old, yet the deposits are at the youngest 11 000 years old and span some 40 000 years.[25]

Despite the remarkable appearance of these remains, morphological preservation often does not correlate with molecular preservation. It was disappointing that repeated attempts at extracting DNA from these coprolites had consistently failed.[26] At that time, Evershed and colleagues[27] had shown that 2000-year-old radish seeds from Egypt contained trapped volatile compounds within their

FIGURE 5. Cross section of strata from Rampart Cave, Arizona, showing some 40 000 years of fossilized ground sloth coprolites. The section is over 2 m deep.

seed coats. These volatiles compounds were products of the Maillard reaction, that is, the reaction of carbonyl groups on reducing sugars with the primary amines of amino acids. This reaction involves many subsequent steps, including a Strecker degradation with multiple condensations, resulting in what is known as "advanced glycosylation endproducts" (AGEs), a largely insoluble, heavily cross-linked mass.[28] Some geochemists believe that these cross-linking reactions and rearrangements make up an integral part of what is collectively termed "humic acids".

To determine whether these coprolites contain similar volatile components, we performed desorption headspace GC/MS on a small section of one coprolite dating to 20 000 years before present (BP). The resultant chromatogram indicated the presence of alkylpyazines, furanones, and furaldehydes, all products of the Maillard reaction (Figure 7). Thus, coprolites stemming from the late Pleistocene extinct animals still contain the volatile components that characterize the Maillard reaction. If volatile components could be retained in the coprolite, then other molecular components should also be present; however, their accessibility with current molecular techniques is questionable.

FIGURE 6. Ground sloth coprolite from Gypsum Cave, Nevada, radio carbon dated at ca. 20 000 years BP. A small incision, where material was removed for headspace GC/MS and DNA analysis, is shown. Scale bar, 10 cm.

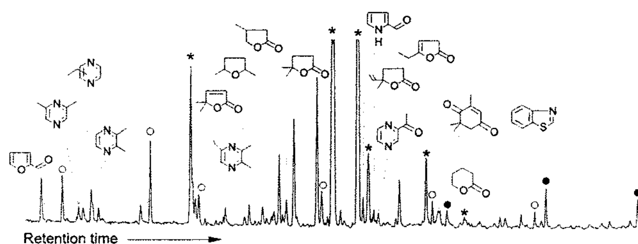

FIGURE 7. Total ion chromatograms of the volatile compounds from a 20 000-year-old ground sloth coprolite from Gypsum Cave, Nevada, analyzed by desorption headspace GC/MS. Chemical structures are shown on figure. These are pyrazines, furanones, and furaldehydes, all derived from the Maillard reaction.

Science moves forward when two different disciplines collide. In a completely unrelated field, the chemical solution to Maillard-induced cross-linking was being sought. Diabetic patients suffer from a variety of misfortunes, mostly stemming from having excess reducing sugars (glucose) present in their blood. These sugars form active cross-links and lead to cataract formation and arterial hardening. In 1996, Vasan and colleagues[29] published a paper describing the use of a novel chemical, a thiazolium salt termed *N*-phenacylthiazolium bromide (PTB), which cleaves the dicarbonyl intermediate in the formation of AGEs. Actively cross-linked collagen could be de-cross-linked with the use of PTB and resume its natural state, both in vitro and in vivo.[29]

When PTB was added to the DNA extractions of our fossil coprolites, small fragments of DNA some 200 bp in length could be amplified via PCR and thus sequenced. In all cases where no PTB was added to the extraction, no endogenous DNA could be retrieved. In small overlap-

FIGURE 8. (A) Climate and vegetation present today at the spring range in Nevada at 2500 m. (B) Gysum Cave, Nevada, summer of 1999.

ping DNA fragments amplified and sequenced multiple times, we managed to obtain some 700 bases of a mitochondrial gene known as the 12SrDNA gene. When this was compared to the DNA from some 1200 vertebrates in a DNA database known as Genbank, the closest match was to a giant ground sloth, Mylodontidae, that we had sequenced in our laboratory a few years earlier.[30] Thus, it appeared that our coprolite stemmed from another family of ground sloths that once lived in the American southwest during the late Pleistocene. These results were recently confirmed when a bone from the sloth, *Nothrotheriops shastensis*, which lived in the southwest was sequenced for the same DNA region and found to be identical to the DNA from the coprolite. Thus, with the help of a de-cross-linking agent, we could positively identify the defecator as the extinct ground sloth *Nothrotheriops shastensis*.

However, these results represented only half of the picture, for the dietary DNA of the plants that had been ingested by the Shasta ground sloth some 20 000 years ago should also be present. To test this, we amplified a small portion of a gene called the ribulose carboxylase oxidase gene (rbcL) which resides in the DNA of the chloroplasts of all plant cells. These products were cloned, and 120 clones were sequenced until the same sequences repeatedly appeared. The sequences were compared to the some 4000 plant sequences at the DNA database Genbank, and the closest matching plant families were identified. Through these means we found some eight plant families that had formed the diet of this sloth. Interestingly, three of these plants, the yucca, agave, and vitis, no longer occur around the cave today but do occur some 800 m higher and 50 km to the northwest. Not surprisingly, the climate around the cave 20 000 years ago was markedly cooler and slightly more humid than it is today.

To complement the information from this coprolite, we chose to analyze simultaneously an additional five coprolites from two other time periods: two from 30 000, one more from 20 000, and two from 11 000 years before present. In these we found the remains of 13 plant families.[31] While many unique plants can be found among the clones sequenced, the majority pointed to an interesting trend. At 30 000 years, a large percentage of the clones appeared to be those of fruits from pine trees, possible pinyon nuts, at 20 000 years yucca and agave along with typical desert biota, and at 11 000 years mostly desert biota. It is thus likely to assume from these data and much more exhaustive morpholigical data that the landscape around and above the cave was riddled with pine and yucca at 30 000 years BP (Figure 8a). At 20 000 years BP, the pine trees had likely retreated to higher elevations, leaving only the yucca and agave and other typical desert plants around the cave. Finally, at 11 000 years BP, at the brink of the sloth's extinction, the yucca and agave had also retreated to higher elvation, and all that was left was the desert biota, similar to what is still present at the cave locality today (Figure 8b). Was this gradual displacement in the vegetation the reason for the ground sloth's demise, or was it the arrival of big game hunters into North America? The answers to questions such as these may lie waiting in the dung.

DNA and Diet from Archaic Native Americans

In the lower Pecos region of southern Texas are several rock shelters, which contain the desiccated remains of

putative human coprolites. One of these shelters, Hinds Cave, was inhabited by prehistoric hunter-gatherers for some 8000 years. Unfortunately, few artifacts can be found in the cave, so little is known about their lives beyond what can be pieced together from the burnt remains of the animals they ate, as well as the macroscopic analyses of the several layers of paloefeces they left behind.

To determine whether these coprolites did, indeed, stem from Native Americans as well as to elucidate aspects of their diet, we extracted DNA from three paleofecal samples that carbon dated to ca. 2500 years BP. Almost all contemporary Native American mitochondrial DNA clusters into what is known as five haplogroups (A, B, C, D, X), that is, five different DNA types distinguishable from other contemporary DNAs. Native American haplotypes are most common today in Eastern Asia, the likely origin of their descendants. To elucidate if these three samples fell into one of the haplogoups, we amplified and sequenced small sections of their mitochondrial DNA. All three samples had unique sequences, thus indicating that they were deposited by three separate individuals. Two of those samples belonged to haplogroup C and one to haplogroup B. Thus, the coprolites stem indeed from three Archaic Native Americans living in Hinds Cave some 2500 years ago.[32]

We found a total of eight different plants based on the DNA from all three coprolites, all of which still occur at the cave locality today. However, to explore the animal diet of these individuals, we amplified a small section of the mitochondrial 12S rDNA gene and sequenced many clones. We found the DNA of pronghorn antelope, bighorn sheep, and cottontail rabbit in two of the three coprolites. Taken alongside the morphological results, which found the presence of small packrats and fish, it is evident that these Archaic Native Americans ate between two and four animals and four and eight different plants over a 1- to 2-day interval.

Coprolites are thus a remarkable source of ancient DNA. Despite lacking a "hardened" exterior such as that of bone and teeth, it is likely that the condensation products derived from the Maillard reaction have "sandwiched" the molecular information within. This entrapment, coupled with a constant burial temperature as in most cave settings, along with the low relative humidity in the American southwest has probably enabled DNA to preserve for many thousands of years.

On the Future of Museum Collections

As mentioned earlier, most animals that lived on Earth have gone extinct; a small proportion of those are luckily housed in museum collections around the world. Unfortunately, very few curators or museums have the time or money to invest in preservational technology. Despite having methods to preserve specimens morphologically intact, these samples will be of little use for current or future DNA techniques when the genetic material available is small DNA fragments. All these prized museum specimens, from which one can retrieve only a few hundred base pairs with tedious work and reproduction, will be of little use and nothing more than academic. One way to alleviate this problem, to move the fossil collections to the forefront of modern science, is to combine the know-how of preservation with current ongoing museum collections. Nature herself has devised unique ways of preserving some of her organisms from extremely hostile environments such as the boiling water baths of Yellowstone National Park and the dry and hot environment of certain deserts.[33] Certain bacteria produce what is known as an endospore, a protective coating to survive extended periods in a hostile environment. Some nematodes and plants produce excessive amounts of the nonreducing sugar, trehalose, which allows the entire organism to completely desiccate, thereby preserving itself in a glassy state, dry and protected until the first rains reappear (Figure 9). This technology needs to be applied to current collections, as simply drying the tissues will cause irreversible damage, and freezing, which for many museums is cost-prohibitive, may be applicable for DNA and RNA but will not help the protein chemists elucidate certain enzymatic functions. As our current biodiversity suffers losses every day, the time to apply the know-how of preservational technology is right now, not in 10 years when some additional thousands of species will have already gone extinct.

FIGURE 9. Desiccation-tolerant plant which can undergo ca. 98% water loss and yet completely rehydrate and survive. (Reprinted with permission from *Current Biology*).

Due to the recent successes in cloning entire organisms such as sheep from cultured cell lines,[34] it has become fashionable to talk, á la movie *Jurassic Park*, about the cloning of extinct organisms. While it is not surprising to long for the extinct mammalian fauna that must have been a dramatic scene in North America 20 000 years ago,

there is from a technical standpoint no hope of cloning these distant ghosts. It would be comparable to cutting the "only" encyclopedia written in an unknown language into single letter sections and attempting to glue it back together.

Future for Ancient DNA

As in most new fields such as the field of ancient DNA, technological developments will open new avenues of research. The future of this field lies in the ability to coordinate interdisciplinary research which combines the talents of geochemists, protein chemists, molecular biologists, archeologists, paleontologists, and evolutionary geneticists. The continued chemical dissection of preservation and a better understanding of the processes which drive fossilization of tissues and hence DNA preservation will enable us to look for DNA in more fossil types, and will push the age of retrieval back farther in time to when large, fantastic animals roamed the planet.

There are several people who contributed immense amounts of time and effort to my pursuit for geologically aged nucleic acids. They deserve at minimum proper acknowledgment. This includes J. Bada, D. Briggs, A. Cooper, G. Eglinton, R. P. Evershed, V. Haynes, M. Kuch, T. Lindahl, P. S. Martin, G. O. Poinar, Jr., K. Sobolik, A. Stankiewicz, and M. Stoneking. Special thanks to D. Poinar for constructive criticism and patience; S. Pääbo for laboratory space, funding, and more; the Max Planck Society, Ludwig Maximillians Universität München, and the Deutsche Forschungsgemeinschaft for support; and two anonymous reviewers for their constructive comments.

References

(1) Raup, D. M. *Extinction: bad genes or bad luck?* W. W. Norton & Co.: New York, NY, 1991.
(2) Higuchi, R.; Bowman, B.; Freiberger, B.; Ryder, O. A.; Wilson, A. C. DNA sequence from the quagga, an extinct member of the horse family. *Nature* **1984**, *312*, 282–284.
(3) Thomas, R. H.; Schaffner, W.; Wilson, A. C.; Paabo, S. *Nature* **1989**, *340*, 465–467.
(4) Pääbo, S. Ancient DNA, extraction, characterization, molecular cloning and enzymatic amplification. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 1939–1943.
(5) Handt, O.; Richards, M.; Trommsdorff, M.; Kilger, C.; Simanainen, J.; Georgiev, O.; Bauer, K.; Stone, A.; Hedges, R.; Schaffner, W.; Utermann, G.; Sykes, B.; Pääbo, S. Molecular genetic analysis of the Tyrolean ice man. *Science* **1994**, *264*, 1775–1778.
(6) Handt, O.; Krings, M.; Ward, R.; Pääbo S. The Retrieval of Ancient Human DNA sequences. *Am. J. Hum. Genet.* **1996**, *59*, 368–376.
(7) Lindahl, T. Instability and decay in the primary structure of DNA. *Nature* **1993**, *362*, 709–715.
(8) Handt, O.; Höss, M.; Krings, M.; Pääbo S. Ancient DNA: methodological challenges. *Experientia* **1994**, *50*, 524–529.
(9) Cooper, A.; Poinar, H. N. Ancient DNA: Do it right or Not at all. *Science* **2000**, *289*, 1139.
(10) Hofreiter, M.; Serre, D.; Poinar, N. H.; Kuch, M.; Pääbo, S. Ancient DNA. *Nat. Genet. Rev.* **2001**, *2*, 353–359.
(11) Krings, M.; Stone, A.; Schmitz, R.; Krainitzki, H.; Stoneking, M.; Pääbo, S. Neandertal DNA sequences and the origin of modern humans. *Cell* **1997**, *90*, 19–30.
(12) Krings, M.; et al. A view of Neandertal genetic diversity. *Nature Genet.* **2000**, *27*, 155–157.
(13) Eglinton, G.; Logan, G. Molecular Preservation. *Philos. Trans. R. Soc. London B* **1991**, *333*, 315–328.
(14) Albert, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, J. D. *Molecular Biology of the Cell*, 3rd ed.; Garland Publishing: New York, NY, 1994.
(15) Shapiro, R. Damage to DNA caused by hydrolysis. In *Chromosome Damage and Repair*; Seeberg, E., Kleppe, K., Eds.; Plenum Publishing Corp.: New York, NY, 1981; pp 3–12.
(16) Greer, S.; Zamenof, S. Studies on depurination of DNA by heat. *J. Mol. Biol.* **1962**, *4*, 123–141.
(17) Poinar, H. N.; Höss, M.; Bada, J. L.; Pääbo, S. Amino acid racemization and the preservation of ancient DNA. *Science* **1996**, *27*, 864–866.
(18) Friedberg, E. C.; Walker, G. C.; Siede, W. *DNA repair and mutagensis*; ASM Press: Washington, D.C., 1995.
(19) Höss, M.; Jaruga, P.; Zastawny, T. H.; Dizdaroglu, M.; Pääbo, S. DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Res.* **1996**, *24* (7), 1304–1307.
(20) Bada, J. L. Racemization of amino acids in nature. *Int. Sci. Rev.* **1982**, *7*, 30–46.
(21) Bada, J. L. Amino acid cosmogeochemistry. *Philos. Trans. R. Soc. London B* **1991**, *333*, 349–358.
(22) Bada, J. L.; Wang, S. X.; Poinar, H. N.; Pääbo, S.; Poinar, G. O., Jr. Amino acid racemization in amber entombed insects: implications for DNA preservation. *Geochim. Cosmochim. Acta* **1994**, *58* (14), 3131–3135.
(23) Stankeiwicz, A. B., van Bergen, P. F., Eds. *Nitrogen Containing Macromolecules in the Bio- and Geosphere*; ACS Symposium Series 707; American Chemical Society: Washington, D.C., 1998.
(24) Poinar, H. N.; Stankiewicz, B. A. Protein preservation and DNA retrieval from ancient tissues *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8426–8431.
(25) Long, A.; Martin, P. S. Death of American ground sloths. *Science* **1974**, *186*, 638–640.
(26) Höss, M.; Dilling, A.; Currant, A.; Pääbo, S. Molecular phylogeny of the extinct ground sloth Mylodon darwinii. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 181–185.
(27) Evershed, R. P.; Bland, H. A.; van Bergen, P. F.; Carter, J. F.; Horton, M. C.; Rowley-Conwy, P. A. Volatile compounds in archeological plant remains and the Maillard reaction during the decay of organic matter. *Science* **1997**, *278*, 43–433.
(28) Mauron, J. The Maillard reaction in food a critical review from the nutritional standpoint. *Prog. Food Nutr.* **1981**, *5*, 5–35.
(29) Vasan, S.; Zhang, X.; Zhang, X.; Kapurmitou, A.; Bernhagen, J.; Teichberg, S.; Basgen, J.; Wagle, D.; Shih, D.; Terlecky, I.; Bucala, R.; Cerami, A.; Egan, J.; Ulrich, P. An agent cleaving glucose derived protein crosslinks. *Nature* **1996**, *382*, 275–278.
(30) Poinar, H. N.; Hofrieter, M.; Spaulding, G. W.; Martin, P. S.; Stankeiwicz, B. A.; Bland, H.; Evershed, R.; Possnert, G.; Pääbo, S. Molecular coproscopy: Dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science* **1998**, *281*, 402–406.
(31) Hofreiter, M.; Poinar, H. N.; Spaulding, G. S.; Martin, P. S.; Possnert, G.; Pääbo, S. A molecular analysis of ground sloth diet through the last glaciation. *Mol. Ecol.* **2000**, *9*, 1975–1984.
(32) Poinar, H. N.; Kuch, M.; Sobolik, K.; Barnes, I.; Stankiewicz B. A.; Spaulding, G.; Bryant V.; Cooper, A.; Pääbo, S. A molecular analyses of the dietary diversity of three Archaic Native Americans *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4317–4322.
(33) Rothschild, L. J.; Mancinelli, R. L. Life in extreme environments. *Nature* **2001**, *409*, 1092–1101.
(34) Campbell, K. H.; McWhir, J.; Ritchie, W. A.; Wilmut, I. *Nature* **1996**, *380*, 64–66.

AR000207X