Opinion TRENDS in Microbiology Vol.13 No.5 May 2005

Geologically ancient DNA: fact or artefact?

Martin B. Hebsgaard[1,2], Matthew J. Phillips[1] and Eske Willerslev[1,2]

[1]Department of Zoology, University of Oxford, South Parks Road, Oxford, UK OX1 3PS
[2]Center for Evolutionary Genetics, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark

Studies continue to report ancient DNA sequences and viable microbial cells that are many millions of years old. In this paper we evaluate some of the most extravagant claims of geologically ancient DNA. We conclude that although exciting, the reports suffer from inadequate experimental setup and insufficient authentication of results. Consequently, it remains doubtful whether amplifiable DNA sequences and viable bacteria can survive over geological timescales. To enhance the credibility of future studies and assist in discarding false-positive results, we propose a rigorous set of authentication criteria for work with geologically ancient DNA.

Introduction

How long can nucleic acids and cells survive in Earth's geosphere? Although the question is fundamental, it is not easily answered. Theoretical considerations and empirical studies suggest maximal DNA survival of 50 thousand (Kyr) to 1 million (Myr) years [1–11]. However, a series of publications claim that ancient DNA (aDNA) from plants, animals and microbes – even viable bacterial cells – can survive in amber, halite, soft tissue and sediments for up to several hundred Myr [12–29]. Implicit in these publications is a belief that nucleic acids can persist over geological timescales (i.e. DNA sequences >1 Myr old). If authentic, these claims have far reaching biological implications. However, aDNA studies involve a high risk of false-positive results and, accordingly, bear a heavy burden of proof [6,9–11,30,31]. In this paper, we evaluate the foundations on which the claims of geologically ancient DNA (gaDNA) depend.

Contamination: the key issue

The central issue in all claims of aDNA and ancient viable microbes is the authentication of positive results. Most fossil remains do not contain any amplifiable endogenous DNA [6], or at best only tiny amounts of degraded molecules (Box 1, Figure 1). Considering this, the use of non-specific media for culturing and the high sensitivity and ability of the polymerase chain reaction (PCR) to amplify single molecules, the risk of contamination is extremely high [9]. Additionally, the transmission of contamination is complex and poorly understood. Contamination can basically be divided into two main types: laboratory-based and sample-based contamination.

Laboratory-based contamination

This covers all types of contamination related to the laboratories where the experiments are conducted. An important but largely unrecognized source of laboratory-based contamination is PCR product carryover. A successful PCR reaction will produce approximately 10^{12} to 10^{15} molecules [32]. Invisible aerosol droplets will be released into the surroundings when PCR tubes are opened or liquids are transferred, each containing 10^5–10^9 molecules [11]. This means that the DNA concentration from a single PCR aerosol is higher than the concentration of putative aDNA in most processed samples (e.g. $\sim 10^5$–10^6 templates per gram of ancient bones [11] and <1 pg environmental DNA per ml of glacial ice [9]). Therefore, without precautions carryover of PCR products will lead to high levels of amplicons rapidly spreading through laboratories, making it easy to obtain false-positive amplification products.

Even if contamination from product carry over can be

Box 1. DNA degradation

The DNA molecule is relatively unstable compared with other cellular components and will degrade with time if not repaired [2]. In general, post-mortem DNA decay must be fast because most fossil remains of even one hundred to a few thousand years old do not contain amplifiable endogenous DNA [6].

Initial post-mortem genomic decay involves cells being dissolved by cellular enzymes. Subsequent cell rupture releases nutrient-rich fluids, which encourage the growth of environmental microorganisms that contribute to further degradation [55]. Nevertheless, in exceptional circumstances, including rapid desiccation, freezing and high salt concentrations, enzymatic and microbial degradation can be significantly reduced or stopped [6]. In such cases, slower but continuous processes, such as hydrolysis, oxidation and cross-linking, will modify the DNA and finally render it irretrievable [6,9–11] (see Figure 1 in the main text). In contrast to metabolically active cells in which genomic damage is effectively repaired through complex enzymatic pathways, DNA damage accumulates in dead or dormant cells (such as bacterial endospores) [55]. Importantly, it is the environmental conditions of preservation rather than the age of the cell that determines the rate of genomic degradation [11]. The key question here is whether any environmental conditions and genomic protection mechanisms can slow down the rate of spontaneous chemical decay to an extent that allows DNA sequences or even cells to survive for million of years.

Corresponding author: Willerslev, E. (ewillerslev@bi.ku.dk).
Available online 26 March 2005

 0966-842X/$ - see front matter doi:10.1016/j.tim.2005.03.010

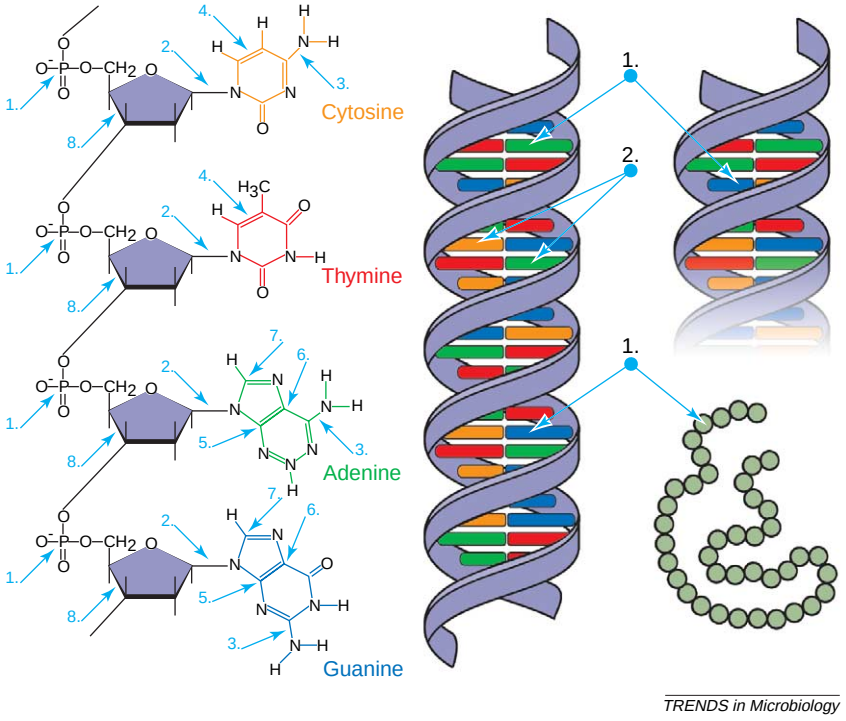

Figure 1. (a) The DNA molecule is highly prone to spontaneous degradation processes, such as hydrolysis and oxidation. Hydrolytic damage is responsible for breaks of the sugar backbone (1.), for base loss (especially the purines, adenine and guanine, referred to as depurination; 2.), and for the deamination of bases (cytosine, adenine and guanine; 3.). Oxidative damage perturbs the integrity of the DNA molecule by attacking the shared double bond of carbons C5 and C6 of pyrimidines (cytosine and thymine; 4.) or the C4 (5.), C5 (6.) and C8 (7.) carbons of purines. The sugar backbone can also be attacked (8.). Hydrolytic and oxidative damage causes nicks and blocking or miscoding lesions. **(b)** A largely unrecognized DNA modification is crosslinking, which includes intermolecular crosslinks (i.e. those of DNA and proteins 1.) and interstrand crosslinks (i.e. between two DNA strands 2.) Crosslinks prevent amplification, but might also stabilize the DNA molecule over time, thereby reducing fragmentation.

minimized, free DNA molecules and cells will be released from laboratory personal, reagents and tools. Even reagents marked sterile can still contain DNA and cells because of moderate sterility assurance levels.

Blank controls cannot by themselves guarantee detection of laboratory contamination because of the sporadic nature of contamination and carrier effects; tiny amounts of contamination will bind to the inside of plasticware and will not be amplified in the blanks. However, adding DNA extract might act as a carrier and release the contaminant molecules for amplification [10,11,30,31].

It is impossible to discount minor amounts of laboratory-based contamination even for the most comprehensive laboratory setup. This holds especially true for human and microbial studies because of the universal distribution of these organisms in laboratory settings [9,10,33]. However, high contamination risk can also be applied to studies of rare organisms (even extinct species) if close modern relatives are processed in the same laboratory or large amounts of amplicons are produced, such as in large-scale population genetic studies [34]. Fortunately, laboratory contamination, although a serious concern, can be addressed by following simple authentication criteria (see below and Box 2).

Sample-based contamination

Here we consider contamination related to the specimen, such as from people handling the specimen during excavation or re-colonisation of ancient materials from the surroundings. Very little is known about levels of sample-based contamination. Recent studies on bones and teeth reveal that sample contamination with human DNA is common and nearly impossible to remove. A cave bear tooth was found to contain 20 different human sequences [6] and multiple human sequences were obtained from archaeological remains even after extensive

Box 2. Summary of criteria needed for authenticating claims of geologically ancient DNA (gaDNA) and cultures

- Specimen well dated to >1 Myr.
- Properly equipped and physically isolated clean laboratory facility.
- Daily movement up the contamination gradient.
- Frequent decontamination of surfaces, reagents and tools in clean laboratory facility.
- Extensive decontamination of ancient specimen before processing.
- Multiple blank controls.
- Spiking PCR with 'non-amplifiable' ancient DNA.
- Amplification of short PCR products (100–500 bp).
- Inverse relationship between amplification length and strength (non-cultures).
- Quantification of template DNA (non-cultures).
- Cloning of PCR products and sequencing of multiple clones.
- Intra-laboratory reproducibility of results (cultures and DNA).
- Independent reproducibility of results by another laboratory (cultures and DNA).
- Passes evolutionary rates tests.
- Ideally, age-dependent pattern in sequence diversity and DNA damage.

Box 3. Predicting long-term DNA survival

Several attempts have been made to predict the long-term survival of DNA, such as amino acid racemization [3], thermal age [7], and extrapolations from DNA in solution [1,9]. In general the models are simplistic, for example, they rely on hydrolytic depurination being the only significant type of DNA damage, whereas other modifications, such as crosslinking, have recently been proven more important for the retrieval of DNA sequences under certain preservation conditions [8,56] (see Figure 1 in the main text). Thus, predicting long-term DNA survival remains problematic, in part because the rates of DNA degradation under various environmental conditions are poorly understood. [57]. Some bacterial cells might even have slow but continuous metabolic activity allowing for some degree of genomic repair over time, although only indirect evidence currently exists for such a scenario [58]. Importantly, claims that long-term DNA survival can be accounted for by alternative and non-measurable metabolic pathways, such as the postulated use of H_2 in spores [25], remain purely speculative and should not be taken into account when predicting DNA survival over geological time spans.

In general, models for long-term DNA preservation predict a maximum survival time of ~100 Kyr for short pieces of amplifiable DNA (~100 bp) [1,3,7]. However, recent studies performed under very strict conditions show that DNA from extinct animals and plants can reproducibly be recovered from frozen sediments dated at 300–400 Kyr, but not from sediments dated to be 1.5–2 Myr [37]. Although these findings could potentially result from leaching of free DNA, many groups currently accept maximum ages for DNA survival of between 100 Kyr and 1 Myr [6–11] on the basis of both theoretical and empirical data.

Despite the problems of modelling the long-term survival of DNA in the geosphere, empirical claims of geologically ancient DNA (gaDNA) in the order of 1000-fold older than theoretical predictions for maximal DNA survival are cause for considerable concern. Together with the huge problems of contamination in aDNA research, this reinforces the burden of proof needed for authenticating claims of multi-million-year-old DNA.

Box 4. Using evolutionary rates tests for authenticity of results

Molecular-distance rates tests and relative rates tests both assume a stochastic molecular clock, for which different lineages accumulate substitutions at the same underlying rate. Molecular-distance rates tests infer how long ago two sequences shared a common ancestor by applying rates estimates to the evolutionary distance between sequences. By contrast, the relative rates test evaluates the null hypothesis that two lineages have evolved at the same rate (see Figure 2 in the main text), or that two lineages (assumed to be evolving at the same rate) are the same age.

Unequal rates of substitution have been attributed to many causes [42], such as differences in metabolic rate, and can lead relative rates tests to either falsely accept as modern or erroneously authenticate a putative geologically ancient DNA (gaDNA) sequence [59]. The relaxed-clock rates test that we have proposed (see Box 5 in the main text) can account for such molecular clock violations. This is because rates can be estimated for all branches on the tree, whereas the rate at the putative gaDNA node can be inferred under alternative constraints for the assumed age of that node. The relaxed-clock rates test also circumvents other limitations of relative rates tests, such as the dependence of statistical power on the distance to the outgroup sequence.

However, if (as for the present case studies) too few related modern sequences are available for generating an expected range of rate variation or reliable fossil calibrations among closely related taxa are not available, the relative rates test is the best alternative.

decontamination by bleach and UV-irradiation [35]. There is no reason to assume that levels of microbial contamination should be less pronounced [9].

In most human and microbial studies, there is currently no way to clearly distinguish an endogenous DNA sequence or culture from that of a contaminant [33]. Novel sequences do not ensure authenticity owing to minor knowledge of, for example, current microbial diversity. Even microbes known to be associated with a particular specimen might have unknown relatives or even identical ecotypes in the surrounding environment [36]. Sample contamination can only be excluded for sequences obtained from morphologically identifiable specimens, with restricted extant distributions and well-known diversity (e.g. many vertebrates and some higher plants), although recent sequencing of DNA directly from sediments [37] complicates authentication for even these groups.

Addressing sample-based contamination

Because it is impossible to totally exclude contamination solely on the basis of experimental setup and sequence identification, it is crucial to apply other tests to infer the authenticity of gaDNA claims. Several rather simplistic approaches have been proposed for predicting the long-term survival of DNA in fossil remains (Box 3). More promising, however, are empirical approaches that exploit the temporal difference between multi-million-year-old gaDNA and modern related sequences.

One such molecular phylogenetic method [38–40] infers the timing of the divergence between gaDNA and modern sequences by assuming a molecular clock and applying a published substitution rate for the particular gene (Box 4). However, this molecular-distance rate test approach can suffer from published rates being unrepresentative for the taxa in question or the portion of the gene used [41]. Furthermore, very old divergences might also be obtained if the putative gaDNA is from a previously unknown modern contaminant.

The relative rates test [38,39] provides a more solid approach, which tests for the distance between the outgroup and the ancient sequence and determines whether it is significantly different from the distance between the outgroup and a modern sequence that is closely related to that of the gaDNA (Box 4, Figure 2). This eliminates dependence on accurate calibration dates and substitution rates.

The above molecular dating and relative rates tests assume that substitutions accumulate in a clock-like manner. This is often not true, with substitution rates varying, even among closely related taxa (Box 4) [42]. If, however, the geological date of the putative gaDNA sequence is well constrained, and reliable fossil calibrations for divergences among related taxa are available, then methods that do not assume a molecular clock (relaxed-clock methods, [43,44]) can be used to evaluate the authenticity of gaDNA claims (Boxes 4 and 5). The aim is to test whether the substitution rate estimated for the putative gaDNA fits the range of rate variation among related taxa, given that the putative gaDNA is either it's claimed age (A) or modern (M).

An inferred substitution rate for the putative gaDNA sequence that fits the expected range when that sequence

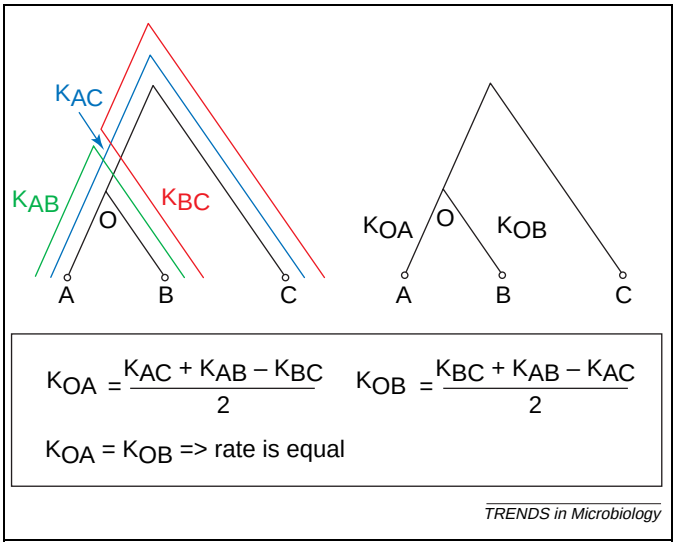

Figure 2. The relative rates test uses an outgroup sequence C, which is known to branch off before either sequence A or B. O is the common ancestor sequence of A and B. K_{OA} is the relative substitution rate between O and A, and K_{OB} is the relative substitution rate between O and B. Whether the genetic distance between O and A is significantly different from the distance between O and B can be evaluated by comparing K_{OA} and K_{OB}, which are calculated using the equations shown.

is fixed at age M, but is substantially higher than the expected range when fixed at age A, would constitute firm evidence that the gaDNA sequence is modern (Figure 3a). Conversely, inferring the substitution rate to be consistent with the expected range when the putative gaDNA sequence is fixed at age A, but substantially lower than this range when fixed at age M is firm evidence against the putative gaDNA sequence being a modern orthologue. We recommend placing confidence intervals on the rate estimates to account for the influence of stochastic error on relaxed-clock rates tests.

Amplification artefacts

Sequence artefacts might bias the results of the evolutionary rates tests (Box 4, Figure 3). Erroneous chimeric DNA sequences can be produced during PCR from two or more template molecules owing to incomplete strand synthesis and blocking lesions (Figure 1). The frequency of chimera production will increase with DNA damage, the number of PCR cycles and the sequence similarity between mixed templates, and can constitute up to 30% of the amplification product [45]. Only those chimeras originating from templates that are sequenced can be detected. Additionally, templates containing stable secondary structures, such as 'hair pins', can yield deletion mutations during PCR [45].

Regular DNA polymerase errors and damage-based miscoding lesions (Figure 1) can result in erroneous point mutations during amplification. The regular polymerase errors can be minimized using high-fidelity polymerase enzymes with proofreading activity, such as *pfu* [46]. Miscoding lesions appear to be more problematic. Recent results suggest they accumulate in hotspots that are distributed similarly to those of evolutionary mutations and hence can cause sequence misidentifications [47].

For the amplification of eukaryotic organellar DNA, such as mitochondrial DNA (mtDNA) and chloroplast DNA (cpDNA) sequences, the adverse amplification of nuclear DNA copies (numts, nuclear mtDNA insertions; and nucps, nuclear cpDNA insertions) remains a threat to sequence authenticity. Upon entering the nuclear DNA (nuDNA), the copy and the original organelle DNA diverge but at different rates, therefore, amplification of numts and nucps can mislead both phylogenetic and evolutionary rate studies. Notably, an early claim of authentic dinosaur DNA obtained from Cretaceous bone fragments [22] was later found to be a human numt [48]. Other forms of paralogy could be similarly misleading.

Molecular-distance rates tests, relative rates tests, and the relaxed-clock rates test approaches are each constrained by several limitations and assumptions (Box 4). However, the latter two methods are currently the most reliable approaches for controlling amplification artefacts. Sequencing artifacts, such as miscoding lesions and chimeric DNA, will typically result in an apparent rate increase on the putative gaDNA branch (Figure 3c). The relative rates and relaxed-clock rates tests distinguish this pattern from that expected for authentic gaDNA (Figure 3b), although the molecular-distance rate test cannot. Unfortunately, there are circumstances in which amplification artefacts can mimic this pattern, such that even the relaxed-clock rates test cannot on its own fully establish the authenticity of gaDNA claims.

Authentication criteria

From the problems outlined in the sections above, it is clear that authenticating gaDNA results is not an easy task, especially if it involves microbes or humans. In fact, the scientific method means that we can never prove the authenticity of a result, we can simply test the result as rigorously as possible and fail to disprove it. This should in no way lessen the importance of authentication testing, especially where results are as susceptible to error as in aDNA research. In Box 2 we have summarized a minimum set of criteria we believe should be fulfilled before a DNA sequence might genuinely be considered geologically ancient. Each criterion is explained in more detail below.

To minimize any laboratory-based contamination the initial culturing experiments or aDNA extraction and PCR setup should be conducted in a dedicated clean laboratory facility (i.e. a laboratory physically isolated from any laboratory where modern day cultures or any PCR products are or have been processed). The clean laboratory must have an isolated ventilation system, nightly UV-irradiation of surfaces and ideally, positive air pressure. Surfaces, including benches, floors, walls and equipment, should be cleaned by 5% bleach or decon on a daily to weekly basis. Personnel must wear full bodysuits, facemasks and gamma-sterilized disposable gloves. Daily movement should always be up the contamination gradient (from the clean laboratory to the post-PCR or culturing laboratory) to minimize product carryover. Reagents, tools and samples need to be cleaned before use, for example, by ultra-filtration (30 and 50 Dalton), UV-irradiation (45W, 72 hours), baking (>180°C, 12 hours), acid (2.5 M HCl, 48 hours) and/or sodium hypochlorite (5%, 48 hours), therefore minimizing contamination from manipulation [11].

Box 5. The relaxed-clock rates test

To illustrate the use of the relaxed-clock rates test, we provide a simulated example based on modern mitochondrial (mt) genome protein-coding sequences from whales (Cetacea) and their relatives. GenBank accession numbers and calibration dates are provided in the supplementary material available online. Maximum-likelihood (within PAUP* 4.0b10, [60]) branch lengths were inferred under a GTR+I+Γ_4 model with the relationships fixed for the whales and other laurasiatherians in accordance with Arnason *et al.* [61] and Murphy *et al.* [62], respectively. Penalized likelihood divergence estimation (within r8s [43]) was used with six calibration bounds to infer substitution rates at every node. The rates at internal and external cetacean nodes provide the expected range for substitution rates among whales (Figure I, grey bars).

The next step was to simulate 100 replicate (using seq-gen 1.3.1 [63]) datasets with substitution parameters and branch lengths as inferred for the original data. However, we included a hypothetical geologically ancient DNA (gaDNA) sequence dated at 12 Myr, constrained to have diverged from the pygmy sperm whale (*Kogia breviceps*) 20 Myr ago and maintained the same substitution rate as *K. breviceps*. Maximum likelihood trees were then inferred from the replicate datasets and the r8s analyses run on these.

Figure Ia shows the results for the combined mt protein-coding simulations with the fossil sequence fixed at 12 Myr. The relaxed-clock rates test correctly rejects the putative gaDNA sequence being from an unknown modern whale sequence, because when the r8s analysis assumes the sequence to be a modern orthologue, the substitution rate 95% confidence interval (CI) falls outside the expected range for whales. Similarly, when the sequence was fixed at the present time for the whole mt protein-coding simulations, Figure Ib shows the test to correctly reject the putative gaDNA sequence being 12 Myr or older.

The results for the smaller cytochrome c oxidase I gene (COI) sequences are shown in Figure Ic,d. In each case, the relaxed-clock rates test is unable to reject either the ancient or modern hypotheses. It follows that to distinguish between authentic gaDNA and modern orthologues, sequences must capture sufficient variability to overcome stochastic effects. Notably, many claims for gaDNA have been based on partial sequences of slowly evolving genes, such as *rbcL* and 18S rRNA.

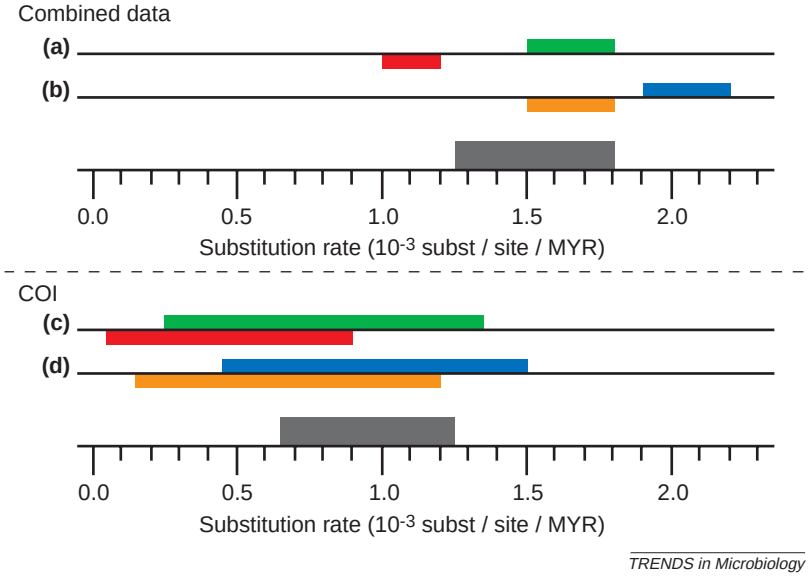

Figure I. 95% confidence intervals (CI) for (penalized likelihood) substitution rate estimates for a hypothetical whale sequence, for which the DNA is (i) 12Myr old and assumed to be 12 Myr old (green); (ii) 12 Myr old and assumed to be modern (red); (iii) modern and assumed to be 12 Myr old (blue); or (iv) modern and assumed to be modern (green). In the relaxed-clock rates test, these intervals are compared with the expected range (grey), which corresponds to the rate variation inferred (under penalized likelihood) for modern whales. The hypothetical whale sequence was constrained to have diverged from the GenBank pygmy sperm whale sequence (20 Myr ago) in trees inferred from data simulated under the maximum-likelihood (GTR+I+Γ_4) model and branch lengths that were estimated for 26 laurasiatherian mammals from **(a,b)** combined mt protein-coding sequences and **(c,d)** the COI gene alone. Although the false assumptions (shown in red and blue) can be rejected for the combined sequences, this is not possible for the smaller and (on average) more conserved COI gene, due to a greater relative influence of stochastic error.

In addition to the ancient sample, blank controls must also be processed and ideally should cover every experimental step. A minimum set of blank controls includes DNA extraction and culture/PCR controls [9]. Positive controls can be useful, but should never be added inside the clean laboratory. It is wise to spike the PCR (including the blank controls) with aDNA that cannot be amplified with the primer sets being used, to test for carrier effects [10,11,30].

Cloning of amplification products should always be performed for the discovery of contamination, PCR artefacts, numts and nucps. Amplifying short PCR products in the 100–500 base pair (bp) size range minimizes the likelihood of PCR artefacts [30]. Quantification of DNA should be performed (e.g. by real-time PCR or other methods) to investigate for the likelihood of amplifying sporadic low-level contamination and for miscoding lesions showing up in the consensus sequences [6,31]. Amplification products should show an inverse relationship between amplification strength and length, owing to the expected fragmentation of the template DNA (Figure 1a). This holds true even when viable cells are present in the sample, because for each viable cell, many more dead cells of the same type should be expected [9].

Independent reproduction of results by another laboratory, to efficiently exclude all types of laboratory-based contamination, is fundamental to every claim of gaDNA and should never be compromised. Reproducibility should be proven by bootstrap testing for diverse sequences, showing that the sequences obtained by independent

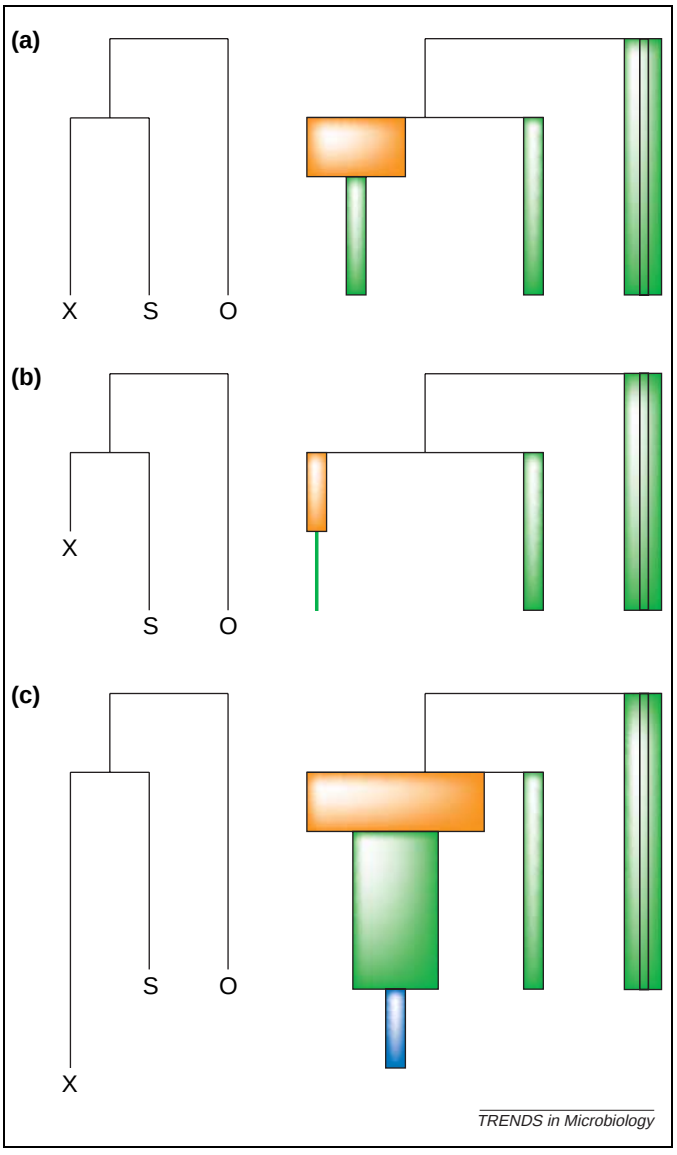

Figure 3. Relaxed-clock rates tests can distinguish between the three basic branch length patterns among the putative geologically ancient DNA (gaDNA) sequence (marked X), its sister taxon (S) and related outgroups (O). The widths of the branches in the right series reflect their estimated substitution rates. These rates depend on the assumed age of X, and will decrease as that age is shifted from the putative geological age (orange), to modern (green) and, theoretically, into the future (blue). The relaxed-clock rates tests rejects rates that do not fall within the expected range, indicated as bounded by the inner and outer widths of the outgroup branch. Hence, we would infer that the putative gaDNA sequence is in fact modern in **(a)**, ancient in **(b)** and alternatively, that the sister and outgroup sequences are ancient in **(c)**. Amplification artefacts that could mimic expectations for gaDNA might comprise a sister sequence that contains errors or that is chimeric. Alternative possibilities for producing this pattern are: (i) chimeras involving a rapidly evolving modern sequence and a distantly related, slowly evolving modern homologue (which together, group with homologues of the faster sequence); and (ii) numts that are homologous to (and evolving more slowly than) poorly conserved mtDNA regions. However, most sequencing errors and chimeras involving the putative gaDNA sequence are expected to increase the branch length of X and therefore produce the spurious result in (c.).

laboratories originate from the same underlying distribution [8,9,37]. For viable cultures, reproducibility should be illustrated by obtaining exact sequence matches from two or more cultures that have been retrieved by independent laboratories and by showing that an additional independent laboratory can obtain at least part of the culture's sequence directly from the sample. If it is impossible to divide a specimen for independent replication, additional specimens should be used and must produce congruent results.

To claim that a DNA sequence obtained from a culture or directly from a fossil or an environmental sample is geologically ancient, the specimen needs to be >1 Myr old. Thus, it is crucial that the specimen is reliably dated, which can be difficult outside of the ^{14}C dating range (~50 Kyr). However, methods such as argon-argon, potassium-argon, fission track, electron-spin resonance and geomagnetic reversal dating are quite well established. Other potential methods relate to *in situ*-produced cosmogenic nuclides, such as berylium (^{10}Be), but these are still very much under development [49,50].

Ideally, several specimens of the same type, preserved under similar conditions but of different ages, should be processed. It will make a strong case if such a sample series shows clear time-dependency in sequence diversity and DNA damage, indicating increased levels of DNA degradation with time [8]. Additionally, it must be implied, by the evolutionary rates tests (Box 4), that the sequence obtained is >1 Myr old.

How solid are current claims of geologically ancient DNA?

Since the invention of the PCR, several reports claim to have obtained gaDNA sequences and viable cells. As will be evident below, none of the reports have fulfilled even the most fundamental of the criteria summarized in Box 2.

In 1990, Golenberg *et al*. [12] claimed to have obtained cpDNA sequences from a 17–20 Myr old Magnolia leaf preserved in waterlogged clay deposits. The claim is exceptional in several aspects. The amplification product obtained was 820 bp, although fossil remains preserved under non-frozen conditions will generally only allow for amplification products of <500 bp in size [6]. The fossil had been in direct contact with water, facilitating fragmentation of the DNA molecule by hydrolytic damage (Box 1, Figure 1a). It has been predicted that at 15°C and neutral pH, fully hydrated DNA will be completely depurinated into fragments <800 bp in ~5 Kyr [1]. The Magnolia experiment was conducted in a laboratory used for amplification of DNA from contemporary plants, thereby facilitating contamination by product carryover. Later attempts to obtain cpDNA sequences from similar 17–20 Myr clay-deposited fossils of *Taxodium*, *Magnolia* and *Persea* have produced conflicting results. Some authors report that only prokaryotic sequences of unknown age can be obtained [51], whereas others, including a recent study, report the retrieval of authentic plant cpDNA in the 700–1500 bp size range [13,29]. However, this work was not conducted in dedicated clean laboratory facilities, nor was it cloned, or confirmed by independent reproduction of results. Although the cpDNA sequences published from the 17–20 Myr plant fossils pass a molecular-distance rate test they do not pass the more rigorous relative rates test (Table 1). Overall, the evidence strongly suggests involvement of contamination (probably PCR related).

Several reports claim to have obtained amber-entombed DNA sequences from a 25–40 Myr bee [14,15], 25–30 Myr old termites [16,18], a 120–135 Myr weevil [17],

Table 1. Summary of geologically ancient DNA (gaDNA) results

Specimen[a]	Source[b]	Age (Myr)[c]	Gene[d]	Size (bp)[e]	Culture (viability)[f]	Sample decont.[g]	Independent replication[h]	Clean room facility[i]	Cloning[j]	Molecular test[k]	Relative rate test[k]	Refs
Amber	Weevil	120–135	18S	315, 226	–	√	–	–	–	Fail	Fail	[17]
Amber	Legume	35–40	*rbcL*	348	–	–	–	–	–	Pass	Fail	[19]
Amber	*Bacillus*	25–40	16S	530	√	√	–	–	–	Pass	Fail	[23]
Amber	*Staphylococcus*	25–35	16S	1548	√	√	–	–	–	Fail	Fail	[24]
Amber	Termite	25–30	16S	94	–	–	–	–	–	Pass	Fail	[16]
Amber	Termite	25–30	18S	239	–	–	–	–	–	Fail	Fail	[16]
Clarckia	Magnolia	17–20	*rbcL*	820	–	–	–	–	–	Pass	Fail	[12]
Clarckia	*Taxodium*	17–20	*rbcL*	1320	–	–	–	–	–	Pass	Fail	[13]
Clarckia	*Persea*	17–20	*rbcL*	699	–	–	–	–	–	Pass	Fail	[29]
Clarckia	*Magnolia*	17–20	*ndhF*	1528	–	–	–	–	–	Pass	Fail	[29]
Halite	*Bacillus*	250	16S	1560	√	√	–	√	–	Fail	Fail	[26]
Halite	Haloarchaeal/ Proteobacteria	11–415	16S	693–1009	–	√	–	√	√	Pass/Fail	Fail	[28]

[a]Type of fossil specimen.
[b]DNA source.
[c]Claimed age of DNA sequence in million years BP.
[d]Gene fragment sequenced.
[e]Length of fragment amplified directly from source or culture.
[f]Viable cells obtained, √ yes, – no.
[g]Attempts made to decontaminate specimen before DNA extraction/culturing, √ yes, – no.
[h]Replication of results by an independent laboratory, √ yes, – no.
[i]Use of clean laboratory facility for pre-PCR, initial culture work, √ yes, – no.
[j]Cloning done, √ yes, – no.
[k]Sequences passing or failing molecular-distance rates test and relative rates test (see Box 4 in the main text, and supplementary material provided).

a 40 Myr legume [19], and a 25–35 Myr old viable endospore [23]. These results are surprising because in amber inclusions even chitin (that is more time-resistant than DNA) has proven to be highly modified by diagenetic processes [52]. A recent large-scale study performed under strict aDNA conditions failed to obtain endogenous DNA from amber insects despite the application of several different extraction and PCR protocols [4,5]. Even attempts to reproducibly amplify endogenous DNA from the same extracts used in the original claim (i.e. from the 25–40 Myr amber preserved bee [14,15]) have failed [53]. Additionally, parts of the reported 120–135 Myr weevil sequence [17] were subsequently shown to be of fungal origin [38] and none of the amber sequences passed relative rates tests [38] (Table 1). Overall, the reports of DNA from amber suffer from not being conducted in dedicated aDNA laboratory settings, from amplification products not being cloned, and in all cases, lack of authentication by independent replication (Table 1).

Several studies have claimed to obtain DNA from multi-million-year-old bacterial endospores and proteobacteria preserved in halite, some of which could be cultured [26–28]. These claims are, from a theoretical perspective, more likely than are those of gaDNA in amber and waterlogged materials because endospores are known to sustain special features facilitating long-term genomic protection, and because *in vitro* experiments [2] suggest that high salt concentrations facilitate long-term DNA preservation.

The oldest claims for any gaDNA are currently made by Vreeland *et al.* [26] and Fish *et al.* [28], who report sequences (up to >1 Kb) obtained from viable bacterial endospores and proteobacteria that had been preserved in halite crystals for up to 250 Myr and 415 Myr, respectively. The experiments were conducted in dedicated clean laboratory facilities. However, the age of the crystals has been questioned [54], and the results have not been independently replicated. Additionally, the sequences from the salt crystals do not pass the relative rates test for authentication [39] (Table 1). Surprisingly, the sequences from Vreeland *et al.* [26] show only 1–3 substitution differences from contemporary bacterial sequences, whereas known mutation rates among related bacteria would have suggested ~59 differences [40]. Intriguingly, a recent study following strict aDNA criteria (including dedicated laboratory facilities, cloning and replication) found that short pieces of DNA (~120 bp) from endospore-formers (such as *Bacillus* and *Clostridium*) could be obtained from frozen permafrost core samples dated at 400–600 Kyr old, but not from samples dated at 1.5–2 Myr and 8.1 Myr old [8]. Additionally, the results showed that DNA from proteobacteria could not be obtained from samples older than 20–30 Kyr. If these results are generally representative for frozen biomes, they bring into serious question the previous claims of multi-million-year-old bacterial DNA from less promising but much older materials.

Concluding remarks

Currently, there are no authentication criteria that can completely exclude all avenues for contamination in studies of aDNA. This holds especially true for studies on ancient human and microbial remains. However, by following strict criteria for authentication, such as those outlined above and summarized in Box 2, the number of false reports can be significantly reduced. It is concerning that all claims published to date on DNA surviving over geological time spans have not followed the most fundamental of these authentication criteria, which unfortunately renders them unreliable. In order for the broader scientific community to seriously consider the possibility of sequencing gaDNA, much more care is needed in future studies.

Acknowledgements

The Wellcome Trust UK supported Eske Willerslev. Matthew J. Phillips was supported by the BBSRC. We would like to thank B. Shapiro, A. Drummond, A. Rambaut, R. Barnett and J-L. Schwenninger for helpful discussions.

Supplementary data

Supplementary data are available with this article and are discussed in Box 5. Supplementary data associated with this article can be found at doi:10.1016/j.tim.2005.03.010.

References

1 Pääbo, S. and Wilson, A.C. (1991) Miocene DNA sequences – A dream come true? *Curr. Biol.* 1, 45–46
2 Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature* 362, 709–715
3 Poinar, H.N. *et al*. (1996) Amino acid racemization and the preservation of ancient DNA. *Science* 272, 864–866
4 Austin, J.J. *et al*. (1997) Problems of reproducibility – does geologically aDNA survive in amber-preserved insects? *Proc. R. Soc. Lond. B. Biol. Sci.* 264, 467–474
5 Austin, J.J. *et al*. (1997) Palaeontology in a molecular world: The search for authentic ancient DNA. *Trends Ecol. Evol.* 12, 303–306
6 Hofreiter, M. *et al*. (2001) Ancient DNA. *Nat. Rev. Genet.* 2, 353–360
7 Smith, C.I. *et al*. (2001) Neanderthal DNA: Not just old but old and cold? *Nature* 410, 771–772
8 Willerslev, E. *et al*. (2004) Long-term persistence of bacterial DNA. *Curr. Biol.* 14, R9–R10
9 Willerslev, E. *et al*. (2004) Isolation of nucleic acids and cultures from fossil ice and permafrost. *Trends Ecol. Evol.* 19, 141–147
10 Pääbo, S. *et al*. (2004) Genetic analyses from ancient DNA. *Annu. Rev. Genet.* 38, 645–679
11 Willerslev, E. and Cooper, A. (2005) Ancient DNA. *Proc. R. Soc. B.* 272, 3–16
12 Golenberg, E.M. *et al*. (1990) Chloroplast DNA sequence from a Miocene *Magnolia* species. *Nature* 344, 656–658
13 Soltis, P.S. *et al*. (1992) An *rbc*L sequence from a Miocene *Taxodium* (bald cypress). *Proc. Natl. Acad. Sci. U. S. A.* 89, 449–451
14 Cano, R.J. *et al*. (1992) Isolation and partial characterisation of DNA from the bee *Proplebeia dominicana* (Apidae: Hymenoptera) in 25-40 million year old amber. *Med. Sci. Res.* 20, 249–251
15 Cano, R.J. *et al*. (1992) Enzymatic amplification and nucleotide sequencing of portions of the 18S rRNA gene of the bee *Proplebeia dominicana* (Apidae: Hymenoptera) isolated from 25-40 million year old Dominican amber. *Med. Sci. Res.* 20, 619–622
16 DeSalle, R. *et al*. (1992) DNA Sequences from a fossil termite in Oligo-Miocene amber and their phylogenetic implications. *Science* 257, 1933–1936
17 Cano, R.J. *et al*. (1993) Enzymatic amplification and nucleotide sequencing of DNA from 120-135 million year old weevil. *Nature* 363, 536–538
18 DeSalle, R. *et al*. (1993) PCR jumping in clones of 30-million-year-old DNA fragments from amber preserved termites (*Mastotermes electrodominicus*). *Experientia* 49, 906–909
19 Poinar, H.N. *et al*. (1993) DNA from an extinct plant. *Nature* 363, 677
20 Desalle, R. (1994) Implications of ancient DNA for phylogenetic studies. *Experientia* 50, 543–550
21 Kennedy, M.J. *et al*. (1994) Preservation records of microorganisms: evidence of the tenacity of life. *Microbiology* 140, 2513–2529
22 Woodward, S.R. *et al*. (1994) DNA sequence from cretaceous period bone fragments. *Science* 266, 1229–1232
23 Cano, R.J. and Borucki, M.K. (1995) Revival and identification of bacterial spores in 25- to 40-million year-old Dominican amber. *Science* 268, 1060–1064
24 Lambert, L.H. *et al*. (1998) *Staphylococcus succinus sp.*, isolated from Dominican amber. *Int. J. Syst. Bacteriol.* 48, 511–518
25 Morita, R.Y. (2000) Is H_2 the universal energy source for long-term survival? *Microb. Ecol.* 38, 307–320
26 Vreeland, R.H. *et al*. (2000) Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. *Nature* 407, 897–900
27 Vreeland, R.H. and Rosenzweig, W.D. (2002) The question of uniqueness of ancient bacteria. *J. Ind. Microbiol. Biotechnol.* 28, 32–41
28 Fish, S.A. *et al*. (2002) Recovery of 16S ribosomal RNA gene fragments from ancient halite. *Nature* 417, 432–436
29 Kim, S. *et al*. (2004) DNA sequences from Miocene fossils: An *ndhF* sequence of *Magnolia Latahensis* (Magnoliaceae) and an *rbcL* sequence of *Persea pseudocarolinensis* (Lauraceae). *Am. J. Bot.* 91, 615–620
30 Cooper, A. and Poinar, H.N. (2001) Ancient DNA: do it right or not at all. *Science* 18, 289
31 Marota, I. and Rollo, F. (2002) Molecular paleontology. *Cell. Mol. Life Sci.* 59, 97–111
32 Kwok, S. and Higuchi, R. (1989) Avoiding false positives with PCR. *Nature* 339, 237–238
33 Rollo, F. and Marota, I. (1999) How microbial ancient DNA, found in association with human remains, can be interpreted. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 354, 111–119
34 Shapiro, B. *et al*. (2004) Rise and fall of the Beringian steppe bison. *Science* 306, 1561–1565
35 Gilbert, M.T.P. *et al*. Biochemical and physical correlates of DNA contamination in archaeological human bones and teeth excavated at Matera, Italy. *J. Arch. Sci.* (in press)
36 Gilbert, M.T.P. *et al*. (2004) Absence of *Yersinia pestis*-specific DNA in human teeth from five European excavations of putative plague victims. *Microbiology* 150, 341–354
37 Willerslev, E. *et al*. (2003) Diverse plant and animal DNA from Holocene and Pleistocene sedimentary records. *Science* 300, 791–795
38 Gutiérrez, G. and Marín, A. (1998) The most ancient DNA recovered from an amber-preserved specimen may not be as ancient as it seems. *Mol. Biol. Evol.* 15, 926–929
39 Graur, D. and Pupko, T. (2001) The Permian Bacterium that isn't. *Mol. Biol. Evol.* 18, 1143–1146
40 Nickle, D.C. *et al*. (2002) Curiously modern DNA for a "250 Million-Year-Old" bacterium. *J. Mol. Evol.* 54, 134–137
41 Gillespie, J.H. (1991) *The Causes of Molecular Evolution*, Oxford University Press
42 Bromham, L. and Penny, D. (2003) The modern molecular clock. *Nat. Rev. Genet.* 4, 216–224
43 Sanderson, M.J. (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19, 101–109
44 Thorne, J.L. and Kishino, H. (2002) Divergence time and evolutionary rate estimation with multilocus data. *Syst. Biol.* 51, 689–702
45 Wintzingerode, F.V. *et al*. (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol. Rev.* 21, 213–229
46 Hansen, A.J. *et al*. (2001) Statistical evidence for miscoding lesions in ancient DNA templates. *Mol. Biol. Evol.* 18, 262–265
47 Gilbert, M.T.P. *et al*. (2003) Distribution patterns of post-mortem damage in human mitochondrial DNA. *Am. J. Hum. Genet.* 72, 32–47
48 Zischler, H. *et al*. (1995) Detecting dinosaur DNA. *Science* 268, 1192–1193
49 Wagner, G.A. (1998) *Age determination of young rocks and artifacts*, Springer Verlag
50 Ludwig, K.R. *et al*. (2000) Geochronology on the paleoanthropological time scale. *Evol. Anth.* 9, 101–110
51 Sidow, A. *et al*. (1991) Bacterial DNA in Clarkia fossils. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 333, 429–433
52 Stankiewicz, B. *et al*. (1998) Chemical preservation of plants and insects in natural resins. *Proc. R. Soc. Lond. B.* 265, 641–647
53 Kimberly, K.O. *et al*. (1997) Ancient DNA from amber fossil bees? *Mol. Biol. Evol.* 14, 1075–1077
54 Hazen, R.M. and Roedder, E. (2001) How old are bacteria from the Permian age? *Nature* 411, 155–156
55 Gilbert, M.T.P. and Hansen, A.J. (2005) Post mortem damage in aDNA: implications and assessing aDNA quality. In *Molecular Markers, PCR, Bioinformatics and Ancient DNA - Technology, Troubleshooting And Applications* (Dorado, G., ed.), Science Publishers
56 Poinar, H.N. *et al*. (1998) Molecular coproscopy: Dungand Diet of the Extinct Ground Sloth *Nothrotheriops shastensis*. *Science* 281, 402–406
57 Nicholson, W.L. *et al*. (2000) Resistance of *Bacillus* endospores to extreme terrestrial and extraterrestrial environments. *Microbiol. Mol. Biol. Rev.* 64, 548–572
58 Rivkina, E.M. *et al*. (2000) Metabolic activity of permafrost bacteria below the freezing point. *Appl. Environ. Microbiol.* 66, 3230–3233

59 Tajima, F. (1993) Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* 135, 599–607
60 Swofford, D.L. (2003) PAUP*. Phylogenetic analysis using parsimony (*and other methods), Version 4. Sinauer Associates
61 Arnason, U. *et al*. (2004) Mitogenomic analyses provide new insights into cetacean origin and evolution. *Gene* 333, 27–34
62 Murphy, W.J. *et al*. (2001) Resolution of the early placental mammal radiation using bayesian phylogenetics. *Science* 294, 2348–2351
63 Rambaut, A. and Grassly, N.C. (1997) Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comput. Appl. Biosci.* 13, 235–238