Ancient bacteria show evidence of DNA repair

Sarah Stewart Johnson*†, Martin B. Hebsgaard‡, Torben R. Christensen‡, Mikhail Mastepanov‡, Rasmus Nielsen‡, Kasper Munch†, Tina Brand‡, M. Thomas P. Gilbert†, Maria T. Zuber*, Michael Bunce§, Regin Rønn†, David Gilichinsky§, Duane Froese, and Eske Willerslev*‡**

*Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, 54-810, Cambridge, MA 02139; †Centre for Ancient Genetics and Centre for Comparative Genomics, Institute of Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark; ‡Geobiosphere Science Centre, Lund University, Sölvegatan 12, Lund, 22362 Sweden; ‡Ancient DNA Research Laboratory, Murdoch University, Perth, WA 6150, Australia; ‡Soil Cryology Laboratory, Institute of Physicochemical and Biological Problems in Soil Science, Russian Academy of Sciences, Pushchino, Russia; and Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, AB, Canada T6G283

Recent claims of cultivable ancient bacteria within sealed environments highlight our limited understanding of the mechanisms behind long-term cell survival. It remains unclear how dormancy, a favored explanation for extended cellular persistence, can cope with spontaneous genomic decay over geological timescales. There has been no direct evidence in ancient microbes for the most likely mechanism, active DNA repair, or for the metabolic activity necessary to sustain it. In this paper, we couple PCR and enzymatic treatment of DNA with direct respiration measurements to investigate long-term survival of bacteria sealed in frozen conditions for up to one million years. Our results show evidence of bacterial survival in samples up to half a million years in age, making this the oldest independently authenticated DNA to date obtained from viable cells. Additionally, we find strong evidence that this long-term survival is closely tied to cellular metabolic activity and DNA repair that over time proves to be superior to dormancy as a mechanism in sustaining bacteria viability.

In recent years, a number of studies have claimed that ancient bacterial cells and their DNA can survive for many millions of years within sediments, amber, and halite (e.g., refs. 1–4). The most common explanation for these findings is that the microbes have remained in a stage of dormancy, known to be associated with high stress tolerance and resistance to adverse conditions. Although dormancy can be followed by special adaptations that reduce the rate of DNA damage, truly dormant cells, like the endospores of *Bacillus* and *Clostridium*, remain metabolically inactive and therefore have no active DNA repair (5). As a result, their genomes will degrade with time because of spontaneous chemical reactions like hydrolysis and oxidation (6) that finally become fatal, preventing the cell from germinating. Models suggest that unrepared genomic DNA will be fragmented into small pieces <100 bp in size or will become severely crosslinked within at most 100,000 to 1 million years (100 Kyr–1 Ma) under optimal frozen conditions and much faster in warmer settings (6–10). Thus, the controversy of viable ancient bacteria is heightened by an absence of convincing evidence for mechanisms by which a cell can withstand damage to DNA and other unstable molecules such as ATP over geological timescales (11–14). Even though there have been speculations and some indirect evidence of respiration in ancient microbes (e.g., refs. 15–21), so far there has been no direct evidence of active DNA repair. In this study, we used a combination of molecular biology techniques and direct measurement of CO2 production from perennially frozen samples to show that dormancy is inferior to low-level metabolic activity with DNA repair as a long-term survival mechanism in ancient bacteria.

Results and Discussion

We investigated samples from permafrost, because these constant subzero temperature environments are considered among the best for long-term microbial and DNA survival (13). Samples were drilled under strict conditions in northeastern Siberia, northwestern Canada, and Antarctica, and they were kept frozen until they were processed for DNA extraction in the laboratory (Table 1). The cores were spiked on the surface with a recognizable contaminant during drilling, as in refs. 11, 13, and 22, allowing us to test for penetration of contamination during collection, transport, and handling.

To ensure that DNA from dead cells was not included in the study, we attempted to amplify only 4-kb bacterial ribosomal DNA fragments from our samples by using universal bacterial primers. Previous studies have shown that fossil remains of dead organisms rarely produce endogenous amplification products longer than 100–500 bp in size (23), and no report has reproducibly generated amplicons >1,042 bp from a dead specimen on ancient timescales (24). The 4-kb amplicon length is both a factor of 4 beyond the longest fragment ever retrieved from the ancient DNA of dead cells and ~20 times longer than ancient DNA fragments recovered from plants and mammals in the same samples (88–230 bp, chemically similar to the DNA of microbes and certainly obtained from dead biomass) (22). Although the successful culturing of microbes from ancient specimens could serve as direct evidence for life, this traditional tactic has been deliberately avoided, because it suffers from two serious constraints. First, <1% of all cells are believed to be culturable using standard methods (25), severely restricting the applicability of the approach. Second, the long-term incubation times necessary for the detection of low-temperature growth greatly increase the risk of contamination (13).

Six samples dating up to 400–600 Kyr yielded 4-kb amplicons of bacterial DNA, whereas no amplification products were obtained from samples dated to 740 Kyr and ~1 Ma, respectively (Table 1). Attempts to amplify 1 and 4 kb of rDNA from higher plant material in the samples failed. To exclude the possibility of false-positive results because of intralaboratory contamination, permafrost subsamples were sent to Murdoch University (Australia), where 4-kb amplifications of bacterial DNA were independently obtained (Table 1).

The successful and reproducible amplification of 4-kb bacterial DNA but not plant DNA suggests that viable bacterial cells are likely to be present in the permafrost core samples. Importantly, decreasing sequence diversity with age of the recovered bacterial DNA further supports the results’ authenticity: this
Table 1. Permafrost samples analyzed for 4-kb amplification products

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Site</th>
<th>Age range</th>
<th>Generated 4-kb amplifications?</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 0K</td>
<td>Kolyma Lowland, Plakhin Jar (ca. 160° 50'E, 68° 40'N), depth up to 0.5 m</td>
<td>Seasonally frozen modern tundra soil</td>
<td>✓</td>
<td>–</td>
</tr>
<tr>
<td>Sample 7K*</td>
<td>Laptev Sea Coast, Cape Bykovskii (129° 30'E, 71° 40'N), depth, 4.8 m</td>
<td>Holocene 5–9 Kyr</td>
<td>✓</td>
<td>9, 11, 22</td>
</tr>
<tr>
<td>Sample 10K*</td>
<td>Kolyma Lowland, Kon’koyava River (158° 28'E, 69° 23'N), depth 4.0 m</td>
<td>Holocene 10.425 ± 0.045 Kyr</td>
<td>✓</td>
<td>9, 11, 22</td>
</tr>
<tr>
<td>Sample 21K</td>
<td>Ledovyi Obryv Exposure, Main River Ice Bluff, Southern Chukotka (171° 11'E, 64° 06'N), depth, 6.0 m</td>
<td>Late Pleistocene 20.900 ± 0.110 Kyr</td>
<td>✓</td>
<td>–</td>
</tr>
<tr>
<td>Sample 25K**</td>
<td>Kolyma Lowland, Chukochia River (156° 59'E, 69° 29'N), depth, 14.8 m</td>
<td>Late Pleistocene 20–30 Kyr</td>
<td>✓</td>
<td>11, 22</td>
</tr>
<tr>
<td>Sample 500K***</td>
<td>Khomus-Yuryakh River (153° 40'E, 70° 05'N), depth, 41.6 m</td>
<td>Middle Pleistocene 400–600 Kyr</td>
<td>✓</td>
<td>9, 11, 22</td>
</tr>
<tr>
<td>Sample 740K†</td>
<td>Dominion Creek, Yukon (138° 36'E, 63° 41'N), depth, 10 m</td>
<td>Middle Pleistocene 740 ± 60 Kyr</td>
<td>✓</td>
<td>26</td>
</tr>
<tr>
<td>Sample 1M*§</td>
<td>Beacon Valley, Antarctica (160° 36'E, 77° 50'S), depth, 14.5 m</td>
<td>≈1 Ma§</td>
<td>✓</td>
<td>5, 24–29</td>
</tr>
</tbody>
</table>

*Both DNA concentration and the frequency of interstrand cross-links were assayed on these samples in (9, 11). Consistent with the DNA degradation undergone in dead cells, DNA concentration decreases with increasing age, whereas the number of interstrand cross-links increases. Additionally, cell counts from these samples changed little with increasing age. Observed cell counts for three other samples between 1.5 and 2 Ma in age (all with no amplifiable DNA) were similar to these younger samples, thus suggesting that bacterial remnants are well preserved over these timescales (11). This is expected, because the DNA molecule is relatively unstable compared to other cellular components and consistent with the idea that few or no new cells arise.

†Metabolic Activity experiments were conducted on these samples.

‡Four-kilobase amplification results were replicated on a duplicate sample of permafrost 500K in the Ancient DNA Research Laboratory, Murdoch University.

§Sample 1M was taken from beneath an 8.1-Ma volcanic ash layer that has been interpreted as a direct air-fall deposit (27). The antiquity of Sample 1M is supported by a number of studies (28–30). It should be noted, however, that a recent investigation has questioned age relations (31), and analyses are ongoing. Nevertheless, researchers at the University of Washington contend that our sample is at least 1 Ma in age.†† Thus, for our study, we have assigned it an age that we think cautious, and that accords with the available data, ≈1 Ma.

pattern has previously been seen in studies of ancient permafrost samples (11) and is unlikely to result from contamination (Fig. 1; see legend to Table 1 and Materials and Methods). Together with near-constant levels of preserved cellular structures with sample age, the result is consistent with the view that ancient permafrost does not sustain a reproductive bacterial community (13).

Ancient viable bacteria may in principle exist in two different states: (i) a dormant state, such as an endospore, which involves no metabolic activity and therefore no active DNA repair; or (ii) a metabolically active state that may allow for some degree of DNA repair. One way to discriminate between these two states is through assessment of relative levels of DNA damage. The DNA molecule is susceptible to many forms of chemical modification (6, 32). One form commonly observed is the hydrolytic deamination of cytosine, generating uracil or its analogs. The subsequent pairing of uracil with adenine during polymerase amplification leads to the observation of characteristic C→T/G→A transitions (33, 34). To identify metabolically active cells, we determined the relative levels of genetic damage by treating aliquots of the DNA extracts with uracil-β-glycosylase (UNG) before amplification of 4-kb rDNA bacterial fragments. UNG breaks the base-ribose bond in uracil (the product of cytosine deamination) and allows only undamaged DNA to be amplified (Fig. 2; see Materials and Methods). Active in vivo systems can repair this damage; in dead or dormant cells (i.e., cells with no measurable metabolic activity), however, uracil residues will be expected to accumulate over time.

Our analyses of UNG-treated sequences revealed varying levels of DNA damage. In the 5- to 30-Kyr age range, low-GC Gram-positive bacteria with the capacity to form dormant endospores accumulated hydrolytic damage at the 99% confi...
dence level (Fisher exact test, $n = 95; P = 0.00008$). No bacteria with a known capacity for dormancy were detected in the 400- to 600-Kyr amplifications. Instead, members of high-GC Gram-positive Actinobacteria largely related to the nonsporeforming Arthrobacter dominated the oldest intact DNA recovered (Fig. 3).

Seeking evidence of the metabolic activity necessary for DNA repair, we directly tested the same frozen samples for respiration in the form of CO₂ production under close to ambient conditions. Using a highly sensitive technique (see Materials and Methods), we found mean rates of 0.142–0.794 μg of CO₂-C/g dry weight per day in samples <600 Kyr but no CO₂ production above background in the 740-Kyr sample or control blanks, which fits with our inability to amplify long DNA amplification products from these samples (Fig. 4).

Our respiration results together with the lack of DNA damage in high-GC Gram-positive bacteria demonstrate evidence for long-term viability, metabolic activity, and DNA repair in ancient microbial cells. Many studies have suggested that dormancy is the most effective survival strategy for bacteria over long time periods (e.g., refs. 1–3, 35); our data indicate that despite short-term robustness, however, dormant bacteria are unlikely to be the most persistent cells over thousand-year timescales in the cold and desiccated conditions represented by our samples. Instead, bacteria with an active DNA repair mechanism are most likely to persevere.

The long-term survival of bacteria within frozen environments provides a range of intriguing possibilities for DNA maintenance and recovery from subsurface environments. This study demonstrates that permafrost may harbor a subset of viable bacteria adapted to past paleoenvironments, some of which might have yet to be described. The long-term DNA survival observed in Actinobacteria warrants further research, because components of these repair pathways could be enlisted for applications requiring maintenance of DNA integrity for extended periods of time. Finally, to the extent that extant life in permafrost and ice on Mars and Jupiter’s moon Europa is thought to be similar to that on Earth, this study calls for further consideration of metabolically active microbes at subzero temperatures in designing life detection strategies.

Materials and Methods

All prePCR work was carried out in dedicated isolated ancient DNA facilities (with separate ventilation systems, nightly UV irradiation of surfaces, and positive air pressure), and the research team adhered to strict protocols (with full bodysuits, facemasks, and gamma-sterilized gloves) (12, 14). Blank-extraction and PCR-amplification controls were incorporated at ratios of 1:5 and 1:1, respectively. Primary analyses were performed in the Ancient DNA Laboratory at the Centre for Ancient Genetics, University of Copenhagen, Copenhagen, Denmark, and replication of the 4-kb PCR analyses was completed in the Ancient DNA Research Laboratory, Murdoch University (Table 1). The results from the independent laboratories showed an overlap of 83% between sequence groups (i.e., sequences that were ≥96% similar, accounting for intraspecies heterogeneity in the 16S rDNA).

Sample Acquisition. Samples were drilled in northwestern Canada, northeastern Siberia, and Antarctica from sections that have remained frozen since deposition (11, 22). The drilling apparatus was spiked with recognizable bacterial cells or vector DNA for detection of contamination during sampling and handling as described in refs. 9, 11, and 22. Two to four centimeters of the contaminated surfaces were removed with a sterilized microtome knife as in ref. 36, and samples were dated by using fission-track dating, tephrochronology, radiocarbon, and argon dating (11, 22, 26, 33). A previous study of the same permafrost cores suggests that no leaching of free DNA or cells has taken place between the strata (9).

DNA Extraction and Amplification. DNA was extracted and purified (from 2 g of wet weight of sediment) by using established protocols (11, 22). The primer pairs used for the 4-kb bacterial
DNA amplifications were: 341F: 5'-CTTCTACGGAG-GCAGCAGTTGGGAATATTGC-3' and 805R: 5'-GGTGGCAAACTTCCGGAAGGAATTCGAC-3', located on the 16S rDNA and 2167F: 5'-GGTGGCAAACTTCCGGAAGGAATTCGAC-3' and 2637R: 5'-GGTGGCAAACTTCCGGAAGGAATTCGAC-3', located on the 23S rDNA. The primer pairs used for 1- and 4-kb amplifications of plant DNA were: PL1000F: 5'-GTGGGCAAGTTGGCCTTGCCACGATCCACTGAG-3' and PL1000R: 5'-CCAAGAATTTCACCCTTCTGACTATGGAATAC-3', located on the 18S rDNA. Sequences were assigned to taxonomic groups using the GenBank under accession numbers EV083531-EV083798. Recombination as in ref. 36. The sequences are deposited in GenBank under accession numbers EV083531-EV083798.

PCR amplifications were performed in 25-μl reaction volumes: 9 μl of GATC mix (20 mM M/0.25 μl of dNTPs + ddH2O2), 2.5 μl of primer 341F, 2.5 μl of primer 2167T, 2.5 μl of MgSO4, 0.2 μl High Fidelity (HiFi) enzyme (Invitrogen, Carlsbad, CA) with 2.5 μl of HiFi buffer, in addition to 4 μl of BSA and 1.75 μl of DMSO (to aid the denaturing of GC-rich strands). PCR conditions for the nonUNG-treated DNA extracts were: 2 min at 92°C initial; 10 cycles (2 min at 94°C, 1 min at 50°C, 3 min 40 sec at 72°C); 40 cycles (2 min at 94°C, 1 min at 50°C, 3 min 40 sec + 20 additional sec/cycle at 72°C); and 10 min 72°C final. For UNG-treated extracts, 0.25 μl of UNG (Nordic BioSite, Taby, Sweden) and 2.5 μl of UNG buffer (10×) were added to the DNA initially and allowed to incubate at 37°C for 30 min. An initial one-time UNG activation step of 5 min at 50°C was added to the above PCR program. The initial denaturation step at 92°C was also lengthened from 2 to 5 min to completely deactivate the enzyme and prompt strand breaks in damaged templates. The efficacy of UNG for this purpose is supported by refs. 32, 37, and 38. It should be noted that UNG preferentially targets C-rich sequences, therefore there is less detection of damage in low-GC Gram-positive bacteria than in high-GC Gram-positive bacteria. The rate of damage in endospore-forming low-GC Gram-positive bacteria may be even higher than reported, which adds further support to our conclusions.

Cloning and Sequencing. From the 4-kb products amplified, 600-bp fragments were cut to enable cloning and sequencing using the following PCR primer pairs: 907F: 5'-AACCCTACAATGGAC-G3' and rP1: 5'-AGCGTTAATCTTGGAC-GACTT-3' (11). One to three amplifications per sample were pooled, cloned, purified, and sequenced on both strands. The resulting sequences were aligned and investigated for possible recombination as in ref. 36. The sequences are deposited in GenBank under accession numbers EV083531-EV083798.

Sequence Identification. Sequences were assigned to taxonomic groups by using a Bayesian assignment criterion. For each sequence, a BLAST search was performed identifying the five sequences with the highest E-score. Sequences without a taxonomic identification in GenBank were not included. The sequences were first aligned by using ClustalW (39) and then analyzed by using MrBayes (40). For each alignment, two runs of 10,000,000 iterations were performed in MrBayes under the default settings. posterior probabilities of monophyly were inferred from 20,000 sampled trees. A sequence was then assigned to a particular taxonomic group if the probability that it was monophyletic with that group exceeded 90%. A phylogenetic identification chart, including hyperlinks to GenBank and sequence distances, can be found online at www.binf.ku.dk/~kasper/taxonomy/bact.respiration.

Metabolic Activity. Past studies attempting to demonstrate viability and metabolic activity in ancient sealed environments have been prone to contamination, relying heavily on culturing, pulverization, or thawing of samples (15–18). For this reason, we used a sensitive low-temperature technique to conduct tests for microbial respiration on undisturbed permafrost cores. Permafrost subsamples from cores 25K, 500K, and 740K (Table 1) were transferred into a cold incubation apparatus and incubated for 9 months at 10°C in a CO2-free atmosphere. Produced CO2 was removed during incubation. The incubations were performed by using a modified version of an experimental technique (41) that reduces the slight possibility of CO2 release from material of plastic (organic) origin. Hence, the incubation chambers and all connecting tubes were made of stainless steel. The CO2 release was measured in a similar way to that described in ref. 40; after an initial discharge of entrapped CO2 for 3 months, the samples showed a constant level of daily CO2 release during 6 months of incubation. Two control samples (without soil) were processed together with the permafrost samples. Respiration levels for controls as well as for Sample 740K were not distinct from zero, whereas Samples 25K and 500K showed significant relative CO2 release.

The incubations were performed in anaerobic conditions, characteristic of subsurface permafrost environments from Siberia and Canada. Although metabolic activity through chemosynthetic pathways cannot be excluded in the older samples, our genetic findings (no viable bacteria in samples older than 600 Kyr) parallel our respiration results (only CO2 production in samples younger than 600 Kyr). Furthermore, it has been demonstrated that Arthrobacter, the most common genus we detected among the high-GC Gram-positive bacteria, is capable of anaerobic metabolism (42). We thank A. Sher, J. Haile, and G. Zazula for help with sample collection; and J. K. Johnson, S. English, C. Carr, M. Finney, G. Ruvkun, J. Stone, F. Hallet, M. Bhatia, L. Bai, E. Brunskill, and J. Lavinsky for assistance in manuscript revision. We also thank GeneTime; the Marie Curie Research Training Network; the Danish Research Council, Carlsberg, Denmark; The Royal Society; and the Wellcome Trust for support of this work.