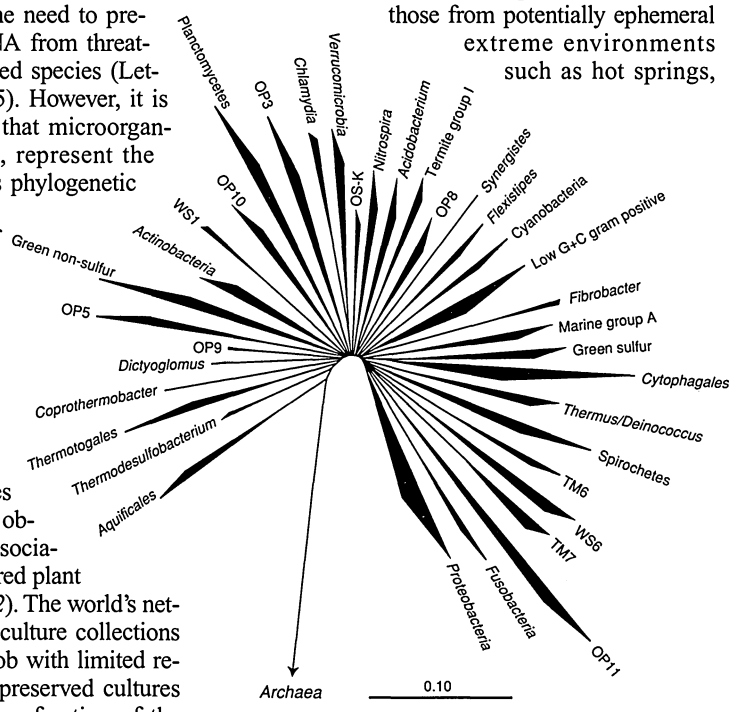
In considering the preservation of species DNA it is noted that "microorganisms, not animals, represent the bulk of the world's phylogenetic diversity and are also in need of preservation in collections." The complexities and triumphs of distinguishing real but very rare or very subtle results from statistical fluctuations with Bayesian analysis are discussed. The benefits and shortcomings of Old World monkeys versus chimpanzees as subjects for analysis in any future primate genome project are examined. And it is observed that the study of artificial genetics has the potential to unite researchers who are on one side or the other of the traditional divide between "natural history" and the "physical sciences."

Microorganisms Should Be High on DNA Preservation List

The Policy Forum "DNA banks for endangered animal species," by Oliver A. Ryder and colleagues (*Science*'s Compass, 14 Apr., p. 275) generated comment from Phillip A. Morin with response from the authors concerning the need to preserve cells and DNA from threatened and endangered species (Letters, 4 Aug., p. 725). However, it is poorly recognized that microorganisms, not animals, represent the bulk of the world's phylogenetic diversity and are also in need of preservation in collections, especially those from habitats under threat of environmental degradation (*1*). Microorganisms may also be considered endangered in cases where they exist in obligate symbiotic associations with endangered plant or animal species (*2*). The world's network of microbial culture collections do a magnificent job with limited resources, but these preserved cultures represent only a tiny fraction of the microbial species present in the environment as most microorganisms are not readily culturable with established techniques (*3*). For example, there are estimated to be at least 36 major lineages (divisions) of the domain *Bacteria* (*4*), of which only four are even moderately well represented in culture collections (see the figure). This limited representation is due to most of new divisions having been recognized by direct molecular methods not involving cultivation. Indeed, many of these divisions do not have even a single representative species preserved in culture collections [e.g., see (*5*)]. Therefore, the concept of ex situ conservation of microbial biodiversity should be extended beyond pure cultures to habitat samples.

Culture collections could be a valuable tool for the task of preserving samples of endangered habitat reference material or DNA derived from it. Samples could include those from potentially ephemeral extreme environments such as hot springs, acid mine drainage sites, and submarine hydrothermal vents, as well as from unique habitats such as rainforest soil, coral reef invertebrate tissue, oligotrophic lake water, and microbial mats in living stromatolites.



Microbial tree of life. Phylogenetic tree of the domain *Bacteria*, based on comparative analysis of 16*S* ribosomal RNA gene sequences, showing currently recognized major divisions. Divisions moderately well represented in culture collections are shown in red. [Adapted from (*4*).] Scale bar indicates changes per nucleotide.

Preservation of habitat reference material could begin with selection of habitat samples that would ensure that all known microbial lineages of life are represented. It would then be possible to preserve reference samples containing at least one representative of all known divisions of *Bacteria* and *Archaea*. This project would be a worthy initiative for support in the International Biodiversity Observation Year 2001–2002 and could be coordinated with existing similar research funded by the National Science Foundation and the National Air and Space Administration, such as Life in Extreme Environments (*6*) and microbial observatories (*7*). The resulting material would provide an invaluable scientific resource for comparison and testing of hypotheses concerning microbial diversity, physiology, and evolution, in addition to storing essential reference material in case of future habitat loss.

John A. Fuerst
Philip Hugenholtz
Department of Microbiology and Parasitology, University of Queensland, Brisbane, Queensland 4072, Australia. E-mail: fuerst@biosci.uq.edu.au and philiph@biosci.uq.edu.au

References

1. N. R. Pace, *Science* **276**, 734 (1997); C. R. Woese, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11043 (1998); A. T. Bull, A. C. Ward, M. Goodfellow, *Microbiol. Mol. Biol. Rev.* **64**, 573 (2000).
2. J. T. Staley, *Curr. Opin. Biotechnol.* **8**, 340 (1997).
3. R. I. Amann, W. Ludwig, K.-H. Schleifer, *Microbiol. Rev.* **59**, 143 (1995).
4. P. Hugenholtz, B. M. Goebel, N. R. Pace, *J. Bacteriol.* **180**, 4765 (1998).
5. M. A. Dojka, J. K. Harris, N. R. Pace, *Appl. Env. Microbiol.* **66**, 1617 (2000).
6. http://www.nsf.gov/pubs/2000/nsf0037/nsf0037.htm
7. http://www.nsf.gov/pubs/2000/nsf0021/nsf0021.htm

Figuring the Odds

Bayesians are not surprised that results at several standard deviations are often spurious, as Charles Seife points out in his News Focus report "CERN's gamble shows perils, rewards of playing the odds" (29 Sept., p. 2260). The significance levels he discusses overstate the evidence, often by a great deal (*1*). Is a tossed coin that gave 60 heads and 40 tails fair? In the sidebar, "A Greek letter, demystified" (p. 2261), it is incorrectly stated that this happens only 2% of the time with a fair coin. In fact, the probability of obtaining exactly 60 heads and 40 tails is 0.011 (from the binomial distribution). What was apparently meant was that the probability of obtaining 60 or more heads out of 100 is 2%, the one-sided *P*-value (but this is actually 2.8%).

The *P*-value is a poor measure of the evidence against "no bias." If a biased coin averages 60% heads, the probability of obtaining 60 heads and 40 tails is 0.081, so on this evidence the odds against "no bias" are at most 0.081/0.011 or 7.5:1, not 35:1 or 50:1. In addition, this computation is done under the extreme assumption that

the actual bias of the coin equals the observed proportion of heads. A fully Bayesian analysis, accounting for the fact that the amount of bias is unknown, might even mildly favor "no bias"!

I was teaching a unit on Bayesian inference when the discovery of a possible planet around PSR 1829-10 mentioned in Seife's article was announced. I noticed that the period claimed for the planet was suspiciously close to a low harmonic of Earth's orbital period. A simple Bayesian analysis for my class (*2*) indicated the truth: it was highly probable that the result was spurious. Let this be a lesson!

William H. Jefferys
Department of Astronomy, University of Texas, Austin, TX 78712, USA. E-mail: bill@astro.as.utexas.edu

References

1. J. O. Berger and M. Delampady, *Stat. Sci.* **2**, 317 (1987).
2. W. H. Jefferys and J. O. Berger, *Am. Sci.* **80**, 64 (1992).

Examining Priorities for a Primate Genome Project

We agree with the premise of Edwin H. McConkey, Agit Varki, and cosignatories that "[a] primate genome project deserves high priority" in the U.S. biomedical research enterprise (Letters, 25 Aug., p. 1295), but we disagree with their selection of the chimpanzee as the focus for these efforts, and with the recommendation to commit resources to genome sequencing at this time.

The choice of which primate to select for this $100 million project should be driven by the goals of the National Institutes of Health's (NIH) mission to promote human health through research. Although chimpanzees have been vital for many advances in biomedical research and continue to serve critical research needs, they are not a commonly used model. With fewer than 1000 federally supported chimpanzees in U.S. research colonies and a moratorium on breeding, the chimpanzee is unlikely to become a widely used animal model. Because it is not practical to use chimpanzees for most research purposes, a focus on chimpanzee genomics would have little impact on future opportunities and progress in biomedical research.

A candidate for genome analysis.

CREDIT: J. ROGERS

The initial plan to sequence the human genome was deferred until a detailed gene map was constructed, and we contend that constructing gene maps of nonhuman primates deserves priority over sequencing, for the same scientific and cost-benefit reasons. Furthermore, the resources should be devoted to those species that are most widely used as primate models for genetic research on human diseases and for which extensive physiological data are available—namely, baboons and rhesus macaques.

Another high priority for a primate genome project would be research on gene expression. To the extent that this research is supported by the NIH, the driving force behind it should be to develop a better understanding of biological